Advanced Direction-of-Arrival Estimation and Beamforming Techniques for Multiple Antenna Systems

Vom Fachbereich 18
Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt
zur Erlangung des akademischen Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Dissertation

von
Dipl.-Ing. Michael Rübsamen
geboren am 6. März 1980 in Hachenburg, Deutschland.

Referent:	Prof. Ph.D. Alex B. Gershman
Korreferent:	Prof. Ph.D. Nikos D. Sidiropoulos
Tag der Einreichung:	11. April 2011
Tag der mündlichen Prüfung:	19. Mai 2011

D 17
Darmstädter Dissertation
2011

Abstract

In this thesis, we develop advanced techniques and concepts for direction-of-arrival (DOA) estimation and beamforming.

We study narrowband high-resolution search-free DOA estimation methods that can be applied in the case of arbitrary array geometries. We derive an asymptotic first-order performance analysis of the popular manifold separation (MS) and interpolated root-MUSIC techniques, which takes into account the finite sample effect as well as manifold approximation errors. Moreover, we propose two rooting-based DOA estimators for arbitrary arrays. It is demonstrated by means of computer simulations that the proposed estimators provide attractive tradeoffs between DOA estimation performance and computational complexity.

We also develop a novel array geometry design for azimuthal DOA estimation. The proposed array design stems from the design of minimum redundancy arrays (MRAs), but the sensors are not required to lie on a uniform grid. The proposed array design facilitates a novel subspace-based DOA estimation technique, which allows estimating the DOAs of more uncorrelated sources than there are sensors, using only second-order statistics of the received data.

Furthermore, we study robust adaptive beamformers for narrowband and broadband signals. In the narrowband case, we show that the popular beamformer based on one-dimensional (1D) covariance fitting leads to inherently non-optimum results in the presence of interferers. To mitigate the detrimental effect of interferers, we extend the 1D covariance fitting approach to multi-dimensional (MD) covariance fitting, modeling the source steering vectors by means of uncertainty sets. The proposed MD covariance fitting approach leads to a non-convex optimization problem. We develop a convex approximation of this problem, which can be solved, for example, by means of the logarithmic barrier method. The complexity required to compute the barrier function and its first- and second-order derivatives is derived. Simulation results show that the proposed beamformer based on MD covariance fitting achieves improved performance as compared to the state-of-the-art narrowband beamformers in scenarios with large sample support.

In the broadband case, we develop two finite impulse response (FIR) beamformers based on worst-case output power minimization, which use different constraints to maintain the desired signal. These constraints strictly limit the sensitivity to signal steering vector estimation errors. Additionally, these constraints lead to an incentive for a low sensitivity if the true signal steering vectors lie within the presumed uncertainty sets. This incentive becomes stronger with increasing signal powers. We study the relation between the proposed FIR beamformers and the norm-bounded broadband minimum variance distortionless response (MVDR) beamformer. Furthermore, we develop the discrete Fourier transform (DFT) beamformer counterparts of the proposed FIR beamformers. The proposed FIR beamformers control the frequency response towards the desired signal only for a finite set of frequencies. Based on the theory of positive trigonometric polynomials, we also develop a modified version of the first proposed FIR beamformer, which avoids the frequency discretization and associated errors. Our simulation results verify that the proposed beamformers are attractive alternatives to the current state-of-the-art broadband beamformers. In particular, the proposed FIR beamformers provide a significantly improved capability to suppress interferers and noise as compared to previous FIR beamformers based on worst-case output power minimization.

Kurzfassung

Diese Dissertation behandelt neue Verfahren und Konzepte für Richtungsschätzung und Beamforming.

Wir untersuchen hochauflösende Richtungsschätzverfahren, die ohne spektrale Suche auskommen und für beliebige Arraygeometrien anwendbar sind. Wir leiten eine Näherung erster Ordnung für die asymptotische Performance der Array Interpolation und Manifold Separation (MS) Schätzer her. Dabei berücksichtigen wir sowohl Fehler aufgrund der endlichen Zahl von Abtastvektoren als auch Approximationsfehler für die Array Mannigfaltigkeit. Desweiteren entwickeln wir zwei neue Richtungsschätzer für beliebige Arraygeometrien. Unsere Simulationsergebnisse zeigen, dass diese Schätzer gute Kompromisse zwischen Schätzgenauigkeit und Rechenkomplexität realisieren.

Außerdem entwickeln wir ein neues Arraygeometrie-Design für azimutale Richtungsschätzung. Das vorgeschlagene Arraygeometrie-Design basiert auf dem Design von Minimum Redundancy Arrays (MRAs). Allerdings ist die Sensoranordnung nicht auf ein kartesisches Gitter beschränkt. Das vorgeschlagene Arraygeometrie-Design ermöglicht ein neues unterraumbasiertes Richtungsschätzverfahren, mit dem die Richtungen von mehr Quellen geschätzt werden können als es Sensoren gibt.

Darüber hinaus untersuchen wir robuste adaptive Beamformer für schmalbandige und breitbandige Signale. Für den weitverbreiteten schmalbandigen Beamformer basierend auf einer eindimensionalen Anpassung an die Kovarianz-Matrix zeigen wir, dass Interferenzen zu inhärent fehlerhaften Ergebnissen führen. Um den nachteiligen Effekt der Interferenzen zu reduzieren, erweitern wir die eindimensionale Anpassung an die Kovarianz-Matrix auf eine mehrdimensionale Anpassung. Dabei modellieren wir die Steuerungsvektoren der Interferenzen mittels Fehlertoleranzmengen. Unser mehrdimensionaler Ansatz führt zu einem nichtkonvexen Optimierungsproblem. Wir leiten eine konvexe Näherung für dieses Problem her, welche z.B. mit dem logarithmischen Barriere Verfahren gelöst werden kann. Wir untersuchen die Rechenkomplexität und zeigen anhand von Simulationen, dass das vorgeschlagene mehrdimensionale Anpassungsverfahren zu deutlichen Performancegewinnen gegenüber den besten bekannten Beamformern führt, wenn genügend Abtastvektoren zur Schätzung der Kovarianz-Matrix zur Verfügung stehen.

Für breitbandige Signale entwickeln wir zwei Finite Impulse Response (FIR) Beamformer basierend auf der Minimierung der Worst-Case Ausgangsleistung. Die vorgeschlagenen Beamformer verwenden unterschiedliche Nebenbedingungen, um die Unterdrückung und Verzerrung des gewünschten Signals zu vermeiden. Wir zeigen, dass diese

Nebenbedingungen die Sensitivität der Beamformer strikt begrenzen. Falls die Signal-Steuerungsvektoren in den dafür angenommenen Fehlertoleranzmengen liegen, führen die Nebenbedingungen der vorgeschlagenen FIR Beamformer zusätzlich zu einem Anreiz für eine geringe Sensitivität. Dieser Anreiz ist umso stärker, je höher die Leistung des gewünschten Signals ist. Wir untersuchen die Beziehung der vorgeschlagenen FIR Beamformer zu dem normbegrenzten Minimum Variance Distortionless Response (MVDR) Beamformer. Zudem entwickeln und untersuchen wir die auf der Diskreten Fourier Transformation (DFT) basierenden Gegenstücke der vorgeschlagenen FIR Beamformer. Die vorgeschlagenen FIR Beamformer kontrollieren die Frequenzantwort für das gewünschte Signal nur für eine diskrete Menge von Frequenzen. Unter Verwendung der Theorie positiver trigonometrischer Polynome modifizieren wir den zuerst vorgeschlagenen FIR Beamformer, so dass die Frequenzdiskretisierung und die damit verbundenen Fehler vermieden werden können. Unsere Simulationsergebnisse zeigen, dass die vorgeschlagenen Beamformer zu den besten bekannten breitbandigen Beamformern gehören. Insbesondere führen die vorgeschlagenen FIR Beamformer zu einer deutlich verbesserten Interferenz- und Rauschunterdrückung als bisher bekannte FIR Beamformer basierend auf der Minimierung der Worst-Case Ausgangsleistung.

Acknowledgments

First and foremost, I thank Professor Alex Gershman for supervising and supporting this thesis. He introduced me to the field of array signal processing and gave me the freedom to explore new ideas. Our numerous discussions strongly influenced this thesis and made my doctoral studies a great pleasure. Many thanks also to my co-supervisor, Professor Nikos Sidiropoulos, for his time and interest in my work.

Moreover, I am indebted to a number of colleagues for valuable discussions and support. Most notably, I am grateful to Professor Marius Pesavento, Professor Sergiy Vorobyov, Dr. Aboulnasr Hassanien, Professor Abdelhak Zoubir, Dr. Marco Moebus, Philipp Heidenreich, and Dr. Ulrich Hammes. Special thanks go to Dr. Chong Meng Samson See for inviting me to Singapore for a research visit at Temasek Labs at Nanyang Technological University.

Above all, I am thankful to Christine Jardak and to my parents for their love and encouragement.

Contents

Chapter 1

Introduction

1.1 Array signal processing

Array signal processing is used in several application areas such as radar, sonar, wireless communications, radio astronomy, seismology, acoustics, and medical imaging [MM80, JD93, God97, Sko01, vT02]. Early contributions to this field have been made mostly in the context of wireless communications and radar systems in the first half of the 20th century. In the second half of the 20th century, the tremendous progress of digital processing hardware led to numerous new developments and applications.

Our work is focussed on coherent receive array processing. The sensors simultaneously measure a spatial field at different locations. The received data depends on the characteristics of the sources, the channels, the noise, and the measurement devices. It is processed to gain information about the sources. Typical objectives are to estimate the number of sources, the source DOAs, locations, velocities, etc. In other cases, the source waveforms or their powers are of primary interest. Due to the large number of applications involving different system models and signal processing objectives, array processing is a broad research field.

In this thesis, we focus on DOA estimation and beamforming, which are two closely related key aspects of array processing. Beamformers are spatial filters used to suppress interferers and noise while the desired signal is maintained. To maintain the desired signal, its spatial signature has to be known. Assuming far-field sources, the source spatial signatures can be estimated by estimating their DOAs. Therefore, beamformers are often applied after the source DOAs have been estimated.

Among the known DOA estimators, the best performance has been achieved by methods based on the maximum likelihood (ML) principle. However, the computational complexity of these estimators is often prohibitively high. Therefore, numerous alternative methods have been proposed, which provide different tradeoffs between DOA estimation performance and computational complexity. Some of these methods are applicable only in the case of specific array geometries. This can be impractical, for example if the sensors have to be mounted on a platform with a certain shape. For this reason, we focus on high-resolution search-free DOA estimation methods, which can be applied to arbitrary arrays. We propose two such estimators and show that they provide attractive tradeoffs between DOA estimation performance and computational complexity.

The array geometry is a fundamental property of any array processing system as it determines the set of locations at which the spatial field is sampled. We propose a novel array geometry design for azimuthal DOA estimation. The proposed array design is related to the MRA principle, but the sensors are not required to lie on a uniform grid. The proposed array design facilitates an associated subspace-based method, which allows estimating the DOAs of more sources than sensors, assuming that the sources are mutually uncorrelated.

In this thesis, we study beamforming with the objective to receive the waveform of the desired signal. A common approach is to minimize the beamformer output power subject to a distortionless constraint for the desired signal. The latter constraint is formulated based on the presumed spatial signature of the desired signal. However, this spatial signature is never known exactly, and estimation errors may lead to a severe undesired suppression of the signal-of-interest, an effect which is commonly referred to as signal self-nulling.

We study narrowband and broadband adaptive beamformers, which are robust against signal self-nulling. Over the last decade, several robust beamformers based on worst-case output power minimization have been proposed. We pay particular attention to these beamformers, and provide novel insights in their robustness. In the narrowband case, we develop a novel beamformer based on MD covariance fitting, which achieves substantial performance improvements as compared to the state-of-the-art narrowband beamformers in scenarios with large sample support.

In the broadband case, we propose two novel FIR beamformers based on worst-case output power minimization, which provide an improved capability to suppress interferers and noise as compared to previous FIR beamformers based on worst-case output power minimization. The relation between the proposed FIR beamformers and the norm-bounded broadband MVDR beamformer is investigated. Moreover, the DFT beamformer counterparts of the proposed FIR beamformers are studied.

The subsequent section describes our contributions in more detail.

1.2 Overview and contributions

In Chapter 2, high-resolution subspace-based search-free DOA estimation methods are discussed, focussing on estimators applicable to arbitrary array geometries. First, the popular interpolated root-MUSIC and MS techniques are reviewed. We derive an asymptotic performance analysis of these techniques, which takes into account the finite sample effect and manifold approximation errors. Subsequently, we propose a root-MUSIC-type DOA estimator based on the DFT of the MUSIC null-spectrum samples. We refer to this approach as Fourier-domain (FD) root-MUSIC. Furthermore, we propose a weighted least squares approximation of the MUSIC null-spectrum function to achieve an improved DOA estimation performance. Our simulation results verify that the proposed techniques are attractive alternatives to the current state-of-the-art DOA estimation methods applicable to arrays of arbitrary geometry.

In Chapter 3, we propose a novel approach to the design of array geometries for azimuthal DOA estimation. The proposed array geometry design is related to the MRA concept, but the array sensors are not required to lie on a uniform grid. Simulation results suggest that the apertures of the proposed arrays increase quadratically with the number of sensors. Based on the proposed array geometry design, we develop a subspace-based DOA estimation technique, which allows estimating the DOAs of more sources than sensors, using only second-order statistics of the received data. This DOA estimation technique is related to the covariance augmentation (CA) technique of [PBNH85], but in contrast to the CA technique, it provides non-ambiguous DOA estimates for the full 360° angular field-of-view.

In Chapter 4, robust adaptive beamforming techniques for narrowband signals are studied. First, the beamformer based on worst-case output power minimization of [VGL03] is reviewed. This beamformer can be formulated equivalently as a 1D covariance fitting problem [LSW03]. We show that the 1D covariance fitting approach leads to inherently non-optimum results in the presence of interferers. To mitigate the detrimental effect of interferers, we extend the 1D covariance fitting approach to MD covariance fitting, taking into account information about the interferer spatial signatures. The MD covariance fitting approach results in refined estimates of the source steering vectors. Subsequently, the beamformer weight vector is computed based on the MVDR concept, using the refined estimate of the signal steering vector. The MD covariance fitting approach leads to a non-convex optimization problem. We develop a local convex approximation of this problem, which belongs to the class of semidefinite programming (SDP) problems. This problem can be solved, for example, by means of the logarithmic barrier method. We show that the barrier function and its first- and second-order derivatives can be computed with the complexity $\mathcal{O}(N^3 + N^2L^2)$, where N is the number of sensors, and L the number of sources. Our simulation results demonstrate that the proposed beamformer based on MD covariance fitting achieves an improved performance as compared to the state-of-the-art robust adaptive beamformers in scenarios with large sample support.

In Chapter 5, robust beamforming techniques for broadband signals are studied. We develop two FIR beamformers based on worst-case output power minimization, which use different constraints to protect the desired signal component. Similar to the constraints of the popular norm-bounded broadband MVDR beamformer, the constraints of the proposed beamformers strictly limit the sensitivity to signal steering vector estimation errors. Additionally, the constraints of the proposed FIR beamformers lead to an incentive for a low sensitivity if the signal steering vectors lie within the presumed uncertainty sets. This incentive becomes stronger with increasing signal powers. A similar incentive does not exist for the norm-bounded broadband MVDR beamformer. We study the relation between the proposed FIR beamformers and the norm-bounded broadband MVDR beamformer. Furthermore, we develop and analyze the DFT beamformer counterparts of the proposed FIR beamformers. The proposed FIR beamformers control the frequency response towards the desired signal only for a finite set of frequencies. Based on the theory of positive trigonometric polynomials, we also develop a modified version of the first proposed FIR beamformer, which avoids the frequency discretization and associated errors. The computational complexities of the proposed FIR beamformers are derived. Our simulation results validate that the proposed beamformers are attractive alternatives to the state-of-the-art broadband beamformers. In particular, the proposed FIR beamformers provide an improved capability to suppress interferers and noise as compared to the previous FIR beamformers based on worst-case output power minimization.

The thesis is based on the following publications:

Book chapters:

- M. Rübsamen, A. El-Keyi, A. B. Gershman, and T. Kirubarajan, "Robust broadband adaptive beamforming using convex optimization," in *Convex Optimization in Signal Processing and Communications*, D. P. Palomar and Y. C. Eldar, Eds. Cambridge University Press, 2010, ch. 9, pp. 315–339.
- M. Rübsamen and A. B. Gershman, "Search-free DOA estimation algorithms for nonuniform sensor arrays," in *Classical and Modern Direction-of-Arrival Estimation*, T. Engin Tuncer and B. Friedlander, Eds. Elsevier, 2009, ch. 5, pp. 161–184.

Journal articles:

- M. Rübsamen and A. B. Gershman, "Robust adaptive beamforming using multi-dimensional covariance fitting," *IEEE Trans. Signal Process.*, submitted.
- M. Rübsamen and A. B. Gershman, "Sparse array design for azimuthal direction-of-arrival estimation," *IEEE Trans. Signal Process.*, submitted.
- A. B. Gershman, M. Rübsamen, and M. Pesavento, "One- and two-dimensional direction-of-arrival estimation: An overview of search-free techniques," *Signal Process.*, vol. 90, no. 5, pp. 1338–1349, 2010.
- M. Rübsamen and A. B. Gershman, "Direction-of-arrival estimation for nonuniform sensor arrays: From manifold separation to Fourier domain MUSIC methods," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 588–599, 2009.

Conference papers:

- M. Rübsamen and A. B. Gershman, "Subspace-based direction-of-arrival estimation for more sources than sensors using planar arrays," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Ma'ale Hahamisha, Israel, Oct. 2010, pp. 21–24.
- M. Rübsamen and A. B. Gershman, "Robust adaptive beamforming based on multi-dimensional covariance fitting," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Dallas, TX, USA, March 2010, pp. 2538–2541.
- M. Rübsamen and A. B. Gershman, "Robust presteered broadband beamforming based on worst-case performance optimization," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Darmstadt, Germany, July 2008, pp. 340–344.
- M. Rübsamen and A. B. Gershman, "Performance analysis of root-MUSIC-based direction-of-arrival estimation for arbitrary non-uniform arrays," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Darmstadt, Germany, March 2008, pp. 381–385.
- M. Rübsamen and A. B. Gershman, "Root-MUSIC based direction-of-arrival estimation methods for arbitrary non-uniform arrays," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Las Vegas, NV, USA, March 2008, pp. 2317–2320.

Chapter 2

Search-free direction-of-arrival estimation for non-uniform arrays

2.1 Background

Over the last three decades, multiple DOA estimation methods have been proposed for specific array geometries such as uniform linear arrays (ULAs), uniform circular arrays (UCAs), or uniform rectangular arrays (URAs). However, in practice it is often not possible to use such arrays, for example because the sensors have to be mounted on a platform with a certain irregular shape.

Among the known DOA estimation methods applicable to arbitrary arrays, the techniques based on the ML principle have been shown to achieve the best mean squared error (MSE) performance [Sch68, WK83, Böh85, Böh86, ZW88]. However, their computational complexity is often prohibitively high, because the likelihood function is non-convex. Finding the global maximum of this function therefore requires a multi-dimensional search.

Several high-resolution DOA estimation methods applicable to arbitrary arrays have been proposed based on the concept of signal and noise subspaces. One of the most popular subspace-based techniques is the multiple signal classification (MUSIC) algorithm [Sch79, BK80]. Compared to the methods based on the ML concept, the MUSIC algorithm leads to a spectral search over a reduced parameter space, and therefore it offers a dramatically lower computational complexity. However, in difficult scenarios, for example in scenarios with weak or closely spaced sources, the performance of the MUSIC algorithm is significantly worse compared to the ML methods.

Other subspace-based DOA estimation methods have been proposed, which provide a significantly improved performance as compared to the traditional spectral MUSIC algorithm. Among the most popular algorithms are root-MUSIC [Bar83, RH89] and ESPRIT (estimation of signal parameters via rotational invariance technique) [PRK86, RPK86]. Since these methods avoid any spectral search, their computational complexity is often lower than that of the spectral MUSIC algorithm. However, root-MUSIC and ESPRIT can be applied only in the case of specific array geometries. In particular, root-MUSIC is applicable only if the sensors are located on a uniform grid, whereas

ESPRIT requires that the array consists of two identical and identically oriented subarrays.

Several extensions of root-MUSIC and ESPRIT to circular and rectangular arrays can be found in [ZM91, MZ94, ZHM96, RMZ99]. Furthermore, root-MUSIC has been generalized to partly-calibrated arrays in [PGW02]. However, the array geometries considered in these papers are still rather specific. In particular, these methods cannot be applied in the case of arbitrary array configurations.

Several approaches have been proposed to extend the concept of root-MUSIC to arbitrary non-uniform arrays (NUAs). For example, the array interpolation method [FW92, Fri93, FW93, WFS95, GB97, SG98, HJO04, HJO05] is based on the approximation of the NUA manifold by a linear function of the manifold of a “virtual” ULA. This allows estimating the DOAs using a modified root-MUSIC technique, which is commonly referred to as interpolated root-MUSIC.

Another approach to extend root-MUSIC to arbitrary NUAs has been reported in [DD94a, DD94b, DD94c]. This approach has been termed manifold separation technique in [BRK07]. The essence of this technique is to approximate each component of the steering vector by its truncated Fourier series. Hence, the NUA manifold can be approximated by a linear function of the Vandermonde vector, which contains in its entries the conventional Fourier basis functions. The Vandermonde structure of the latter vector can be exploited to estimate the source DOAs by means of polynomial rooting.

In this chapter, we present a first-order asymptotic performance analysis of the MS and interpolated root-MUSIC techniques based on the approach of [RH89]. Our performance analysis is new in that it takes into account both the manifold approximation errors and the finite sample effect. This performance analysis shows that within the first-order approximation, the MS and interpolated root-MUSIC techniques lead to the same MSE performance if the manifold approximation errors are small. In this case, the MS and interpolated root-MUSIC techniques achieve similar MSE performance, even though these techniques estimate the DOAs by means of rooting different polynomials.

Then, we propose an alternative rooting-based DOA estimation technique for arbitrary NUAs [RG08c, RG09a, RG09b, GRP10]. We refer to this technique as FD root-MUSIC, as it is based on the DFT of the MUSIC null-spectrum samples. The FD root-MUSIC technique is closely related to the MS technique, but our simulation results show that it provides an improved performance in the asymptotic domain at no additional cost.

A further refinement of this technique is the weighted Fourier-domain (WFD) root-MUSIC technique, which uses a weighted least squares approximation of the MUSIC null-spectrum function to obtain improved DOA estimates.

Our simulation results demonstrate that the proposed methods are attractive alternatives to the MS and interpolated root-MUSIC techniques.

2.1.1 System model

Let us consider an array of N omnidirectional sensors and L ($L < N$) narrowband far-field point sources. The sources are assumed to lie in the xy-plane of the coordinate system. The baseband snapshot vector at the kth time instant is modeled as [vT02]

$$\boldsymbol{x}(k) = \boldsymbol{A}(\boldsymbol{\varphi})\boldsymbol{s}(k) + \boldsymbol{n}(k) \in \mathbb{C}^{N \times 1}, \tag{2.1}$$

where

$$\boldsymbol{A}(\boldsymbol{\varphi}) = [\boldsymbol{a}(\varphi_1), \dots, \boldsymbol{a}(\varphi_L)] \in \mathbb{C}^{N \times L} \tag{2.2}$$

is the steering matrix,

$$\boldsymbol{\varphi} = [\varphi_1, \dots, \varphi_L]^T \in \mathbb{R}^{L \times 1} \tag{2.3}$$

is the vector of source azimuth angles, $\varphi_l \in [-\pi, \pi)$ is the azimuth angle of the lth source,

$$\boldsymbol{a}(\varphi_l) = \left[e^{j\frac{2\pi}{\lambda}(p_{x_1} \cos\varphi_l + p_{y_1} \sin\varphi_l)}, \dots, e^{j\frac{2\pi}{\lambda}(p_{x_N} \cos\varphi_l + p_{y_N} \sin\varphi_l)} \right]^T \in \mathbb{C}^{N \times 1} \tag{2.4}$$

is the steering vector (spatial signature) of the lth source, (p_{x_n}, p_{y_n}) are the xy-coordinates of the nth sensor, λ is the wavelength,

$$\boldsymbol{s}(k) = [s_1(k), \dots, s_L(k)]^T \in \mathbb{C}^{L \times 1} \tag{2.5}$$

is the source waveform vector, $s_l(k)$ is the complex envelope of the lth source, $\boldsymbol{n}(k) \in \mathbb{C}^{N \times 1}$ is the noise vector, and $(\cdot)^T$ denotes the transpose. Figure 2.1 depicts the coordinate system. The baseband signal and noise waveforms are modeled as unknown stationary zero-mean random processes. The noise waveforms are assumed to be mutually uncorrelated and uncorrelated from the signal waveforms.

The array covariance matrix can be expressed as [vT02]

$$\boldsymbol{R}_{\boldsymbol{x}} = \mathrm{E}\{\boldsymbol{x}(k)\boldsymbol{x}^H(k)\} = \boldsymbol{A}(\boldsymbol{\varphi})\boldsymbol{P}\boldsymbol{A}^H(\boldsymbol{\varphi}) + \sigma \boldsymbol{I}_N, \tag{2.6}$$

where

$$\boldsymbol{P} = \mathrm{E}\{\boldsymbol{s}(k)\boldsymbol{s}^H(k)\} \tag{2.7}$$

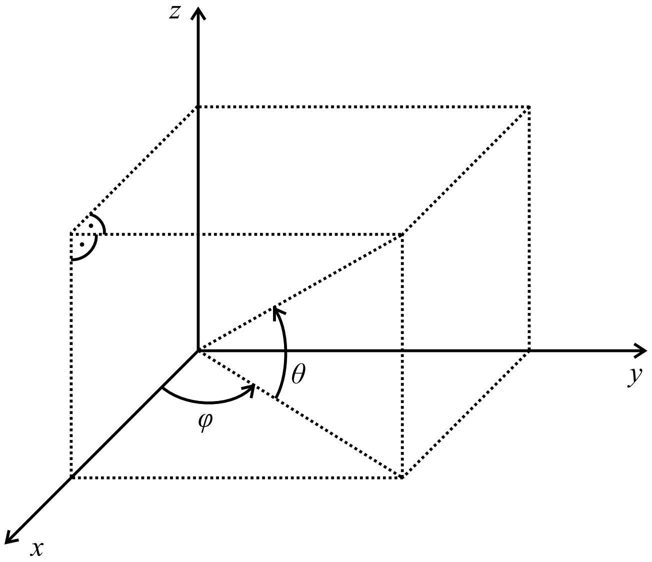


Figure 2.1. Coordinate system.

is the $L \times L$ source waveform covariance matrix, $\mathrm{E}\{\cdot\}$ denotes the statistical expectation, σ is the sensor noise power, $\boldsymbol{I}_N$ stands for the $N \times N$ identity matrix, and $(\cdot)^H$ denotes the Hermitian transpose. The source waveforms are assumed to be non-coherent, so that $\boldsymbol{P}$ is positive definite.

The array covariance matrix can be estimated from the received snapshot vectors as

$$\hat{\boldsymbol{R}}_{\boldsymbol{x}} = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{x}(k)\boldsymbol{x}^H(k), \tag{2.8}$$

where K is the number of snapshots.

2.1.2 Spectral MUSIC

The eigendecomposition of the array covariance matrix can be written as [vT02]

$$\boldsymbol{R}_{\boldsymbol{x}} = \sum_{n=1}^{N} \lambda_{\boldsymbol{x},n}\boldsymbol{v}_{\boldsymbol{x},n}\boldsymbol{v}_{\boldsymbol{x},n}^H, \tag{2.9}$$

where $\lambda_{\boldsymbol{x},n}$ and $\boldsymbol{v}_{\boldsymbol{x},n}$ are the eigenvalues and corresponding orthonormal eigenvectors. Without loss of generality (WLOG), we assume that the eigenvalues $\lambda_{\boldsymbol{x},n}$ are sorted in non-ascending order. Suppose the source steering vectors are linearly independent. Then, the columns of

$$\boldsymbol{V}_{\boldsymbol{x},S} = [\boldsymbol{v}_{\boldsymbol{x},1}, \dots, \boldsymbol{v}_{\boldsymbol{x},L}] \tag{2.10}$$

span the same subspace as the columns of $\boldsymbol{A}(\boldsymbol{\varphi})$, which is commonly referred to as the signal subspace. The orthogonal complement of this subspace, spanned by the columns of

$$\boldsymbol{V}_{\boldsymbol{x},N} = [\boldsymbol{v}_{\boldsymbol{x},L+1}, \dots, \boldsymbol{v}_{\boldsymbol{x},N}], \tag{2.11}$$

is the so-called noise subspace.

Similarly, the eigendecomposition of the sample covariance matrix (2.8) yields

$$\hat{\boldsymbol{R}}_{\boldsymbol{x}} = \sum_{n=1}^{N} \hat{\lambda}_{\boldsymbol{x},n} \hat{\boldsymbol{v}}_{\boldsymbol{x},n} \hat{\boldsymbol{v}}_{\boldsymbol{x},n}^{H}, \tag{2.12}$$

where we assume again that the eigenvalues are sorted in non-ascending order. The estimated signal and noise subspace eigenvector matrices are

$$\begin{aligned} \hat{\boldsymbol{V}}_{\boldsymbol{x},S} &= [\hat{\boldsymbol{v}}_{\boldsymbol{x},1}, \dots, \hat{\boldsymbol{v}}_{\boldsymbol{x},L}] \\ \hat{\boldsymbol{V}}_{\boldsymbol{x},N} &= [\hat{\boldsymbol{v}}_{\boldsymbol{x},L+1}, \dots, \hat{\boldsymbol{v}}_{\boldsymbol{x},N}], \end{aligned} \tag{2.13}$$

respectively.

The MUSIC algorithm exploits that the source steering vectors are orthogonal to the noise subspace eigenvectors, i.e.,

$$\boldsymbol{V}_{\boldsymbol{x},N}^{H} \boldsymbol{a}(\varphi_l) = \boldsymbol{0}, \quad l = 1, \dots, L. \tag{2.14}$$

If

$$\boldsymbol{V}_{\boldsymbol{x},N}^{H} \boldsymbol{a}(\varphi) \neq \boldsymbol{0} \quad \text{for } \varphi \notin \{\varphi_1, \dots, \varphi_L\}, \tag{2.15}$$

then the source DOAs can be estimated by searching for the L smallest minima of the MUSIC null-spectrum function [Sch79, BK80]

$$f_{\mathrm{MU}}(\varphi) \triangleq \boldsymbol{a}^{H}(\varphi) \hat{\boldsymbol{V}}_{\boldsymbol{x},N} \hat{\boldsymbol{V}}_{\boldsymbol{x},N}^{H} \boldsymbol{a}(\varphi). \tag{2.16}$$

2.1.3 Root-MUSIC

The spectral search for the minima of the MUSIC null-spectrum function can be replaced by polynomial rooting if the sensors are located on a uniform grid. For example, in the case of a ULA, the steering vector (2.4) can be written as

$$\boldsymbol{a}(\varphi) = a_1(\varphi) \left[1, e^{j\frac{2\pi}{\lambda}(d_x \cos\varphi + d_y \sin\varphi)}, \dots, e^{j(N-1)\frac{2\pi}{\lambda}(d_x \cos\varphi + d_y \sin\varphi)}\right]^T, \tag{2.17}$$

where (d_x, d_y) is the displacement between adjacent sensors.

Then, the MUSIC null-spectrum function can be expressed as

$$f_{\mathrm{MU}}(\varphi) = \boldsymbol{a}^H(\varphi)\hat{\boldsymbol{V}}_{\boldsymbol{x},N}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}^H\boldsymbol{a}(\varphi) = \boldsymbol{a}_z^T(1/z)\hat{\boldsymbol{V}}_{\boldsymbol{x},N}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}^H\boldsymbol{a}_z(z), \tag{2.18}$$

where

$$z = e^{j\frac{2\pi}{\lambda}(d_x\cos\varphi + d_y\sin\varphi)}, \tag{2.19}$$

and

$$\boldsymbol{a}_z(z) = \left[1, z, \ldots, z^{N-1}\right]^T. \tag{2.20}$$

The root-MUSIC polynomial [Bar83]

$$f_{\mathrm{MU},z}(z) \triangleq z^{N-1}\boldsymbol{a}_z^T(1/z)\hat{\boldsymbol{V}}_{\boldsymbol{x},N}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}^H\boldsymbol{a}_z(z) \tag{2.21}$$

has $2(N-1)$ roots, which form conjugate reciprocal pairs. That means, if z_0 is a root of $f_{\mathrm{MU},z}(z)$, then $1/z_0^*$ is a root as well, where $(\cdot)^*$ denotes the complex conjugate operator. In the noise-free case, $f_{\mathrm{MU},z}(z)$ has pairs of roots at

$$z_l = e^{j\frac{2\pi}{\lambda}(d_x\cos\varphi_l + d_y\sin\varphi_l)}, \quad l = 1, \ldots, L, \tag{2.22}$$

and there are $2(N-L-1)$ additional spurious roots. In the presence of noise, the root locations are distorted, but the signal DOAs can be estimated from the roots of $f_{\mathrm{MU},z}(z)$ that lie closest to the unit circle. Due to the root conjugate reciprocity property, the roots inside the unit circle contain all information about the signal DOAs. Hence, the root-MUSIC algorithm computes all roots of $f_{\mathrm{MU},z}(z)$, and estimates the signal DOAs from the L largest-magnitude roots inside the unit circle.

It has been shown that the root-MUSIC algorithm provides a significantly improved performance compared to the spectral MUSIC algorithm in difficult scenarios, for example if there are weak or closely spaced sources, or if only few snapshot vectors are available to estimate the array covariance matrix. The improved performance of root-MUSIC can be explained by the fact that the DOA estimates are insensitive to radial root errors [RH89].

The root-MUSIC algorithm can be generalized straightforwardly to linear arrays whose sensors lie on a uniform grid. This enables application of root-MUSIC to sparse linear arrays.

2.1.4 Array interpolation

The essence of the array interpolation technique [FW92, Fri93, FW93, WFS95] is to approximate the steering vector of the actual sensor array as

$$\boldsymbol{a}(\varphi) \approx \boldsymbol{G}_{\mathrm{AI}}\boldsymbol{g}_{\mathrm{AI}}(\varphi), \tag{2.23}$$

where $\boldsymbol{g}_{\mathrm{AI}}(\varphi)$ is the $M_{\mathrm{AI}} \times 1$ steering vector of a virtual ULA, M_{AI} is the number of virtual sensors, and $\boldsymbol{G}_{\mathrm{AI}}$ is the $N \times M_{\mathrm{AI}}$ array interpolation matrix. This matrix is chosen so that the approximation error is minimized.

Consequently, the MUSIC null-spectrum function (2.16) can be approximated as

$$f_{\mathrm{MU}}(\varphi) \approx \boldsymbol{g}_{\mathrm{AI}}^{H}(\varphi)\boldsymbol{G}_{\mathrm{AI}}^{H}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}^{H}\boldsymbol{G}_{\mathrm{AI}}\boldsymbol{g}_{\mathrm{AI}}(\varphi) \triangleq f_{\mathrm{AI}}(\varphi). \tag{2.24}$$

Let the displacement between adjacent virtual sensors be (d_x, d_y). Then,

$$f_{\mathrm{AI}}(\varphi) = \boldsymbol{g}_{\mathrm{AI},z}^{T}(1/z)\boldsymbol{G}_{\mathrm{AI}}^{H}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}^{H}\boldsymbol{G}_{\mathrm{AI}}\boldsymbol{g}_{\mathrm{AI},z}(z), \tag{2.25}$$

where

$$\boldsymbol{g}_{\mathrm{AI},z}(z) = \left[1, z, \ldots, z^{M_{\mathrm{AI}}-1}\right]^{T}, \tag{2.26}$$

and z is given by (2.19).

The polynomial

$$f_{\mathrm{AI},z}(z) \triangleq z^{M_{\mathrm{AI}}-1}\boldsymbol{g}_{\mathrm{AI},z}^{T}(1/z)\boldsymbol{G}_{\mathrm{AI}}^{H}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}^{H}\boldsymbol{G}_{\mathrm{AI}}\boldsymbol{g}_{\mathrm{AI},z}(z) \tag{2.27}$$

has $2(M_{\mathrm{AI}} - 1)$ roots, which appear in conjugate reciprocal pairs. Therefore, similar to root-MUSIC, the signal DOAs can be estimated from the largest-magnitude roots located inside the unit circle.

The accuracy of the approximation in (2.23) is typically not sufficient for the whole array angular field-of-view. Therefore, angular sectors have to be defined and such an approximation has to be used separately for each sector [Fri93]. The number of sectors, the number of virtual sensors and their locations are design parameters of the interpolated root-MUSIC method.

2.1.5 Manifold separation

Another root-MUSIC-type technique for arbitrary NUAs is the MS technique [DD94a, DD94b, DD94c]. The essence of this technique is to approximate the steering vector components by their truncated Fourier series expansions. This can be expressed as

$$\boldsymbol{a}(\varphi) \approx \boldsymbol{G}_{\mathrm{MS}}\boldsymbol{g}_{\mathrm{MS}}(\varphi), \tag{2.28}$$

where

$$\boldsymbol{g}_{\mathrm{MS}}(\varphi) = \left[e^{-j\frac{M_{\mathrm{MS}}-1}{2}\varphi}, e^{-j\frac{M_{\mathrm{MS}}-3}{2}\varphi}, \ldots, e^{j\frac{M_{\mathrm{MS}}-1}{2}\varphi}\right]^{T} \tag{2.29}$$

is an $M_{\mathrm{MS}} \times 1$ Vandermonde vector, $\boldsymbol{G}_{\mathrm{MS}}$ is an $N \times M_{\mathrm{MS}}$ matrix, and M_{MS} is assumed to be odd. The nth row of $\boldsymbol{G}_{\mathrm{MS}}$ contains the Fourier series coefficients of the nth steering vector component.

Using (2.28), the MUSIC null-spectrum function can be approximated as

$$f_{\mathrm{MU}}(\varphi) \approx \boldsymbol{g}_{\mathrm{MS}}^{H}(\varphi)\boldsymbol{G}_{\mathrm{MS}}^{H}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}^{H}\boldsymbol{G}_{\mathrm{MS}}\boldsymbol{g}_{\mathrm{MS}}(\varphi) \triangleq f_{\mathrm{MS}}(\varphi). \tag{2.30}$$

The latter function can be expressed in terms of

$$\mathsf{z} = e^{j\varphi} \tag{2.31}$$

as

$$f_{\mathrm{MS}}(\varphi) = \boldsymbol{g}_{\mathrm{MS},\mathsf{z}}^{T}(1/\mathsf{z})\boldsymbol{G}_{\mathrm{MS}}^{H}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}^{H}\boldsymbol{G}_{\mathrm{MS}}\boldsymbol{g}_{\mathrm{MS},\mathsf{z}}(\mathsf{z}), \tag{2.32}$$

where

$$\boldsymbol{g}_{\mathrm{MS},\mathsf{z}}(\mathsf{z}) = \left[1, \mathsf{z}, \ldots, \mathsf{z}^{M_{\mathrm{MS}}-1}\right]^{T}. \tag{2.33}$$

The polynomial

$$f_{\mathrm{MS},\mathsf{z}}(\mathsf{z}) \triangleq \mathsf{z}^{M_{\mathrm{MS}}-1}\boldsymbol{g}_{\mathrm{MS},\mathsf{z}}^{T}(1/\mathsf{z})\boldsymbol{G}_{\mathrm{MS}}^{H}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}\hat{\boldsymbol{V}}_{\boldsymbol{x},N}^{H}\boldsymbol{G}_{\mathrm{MS}}\boldsymbol{g}_{\mathrm{MS},\mathsf{z}}(\mathsf{z}) \tag{2.34}$$

has $2(M_{\mathrm{MS}}-1)$ roots, which appear in conjugate reciprocal pairs [DD94b]. The source DOAs can be estimated from the largest-magnitude roots located inside the unit circle in a way similar to that used in the conventional root-MUSIC algorithm.

The approximation error in (2.28) can be made arbitrarily small by increasing M_{MS}. The MS technique therefore allows estimating the source DOAs by rooting a single polynomial, whereas the array interpolation technique requires computing the roots of several polynomials for different angular sectors. However, the polynomial of the MS technique typically has a larger degree than the polynomials of the array interpolation technique.

To quantify the required value of M_{MS} for the accurate approximation of the array manifold, let us consider the Fourier series expansion of the nth component of the steering vector, that is,

$$a_n(\varphi) = \sum_{m=-\infty}^{\infty} F_{a_n}[m] e^{jm\varphi}, \tag{2.35}$$

where

$$F_{a_n}[m] \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} a_n(\varphi)\, e^{-jm\varphi}\,\mathrm{d}\varphi \tag{2.36}$$

is the mth Fourier series coefficient. Truncating the Fourier series to M_{MS} terms yields

$$a_n(\varphi) \approx \sum_{m=-\frac{M_{\mathrm{MS}}-1}{2}}^{\frac{M_{\mathrm{MS}}-1}{2}} F_{a_n}[m] e^{jm\varphi} \triangleq S_{\frac{M_{\mathrm{MS}}-1}{2}}\{a_n\}(\varphi), \tag{2.37}$$

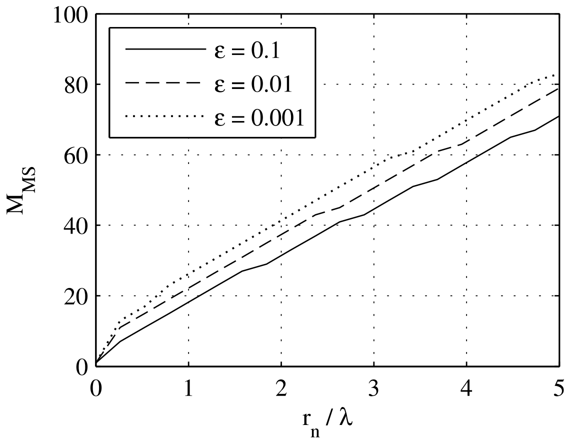


Figure 2.2. Minimum of M_{MS} subject to the constraint (2.39).

where M_{MS} is assumed to be odd. The approximation error can be measured in terms of the L^2-norm as

$$\left\| a_n(\varphi) - S_{\frac{M_{\mathrm{MS}}-1}{2}}\{a_n\}(\varphi) \right\|_{L^2} \triangleq \left(\int_{-\pi}^{\pi} \left| a_n(\varphi) - S_{\frac{M_{\mathrm{MS}}-1}{2}}\{a_n\}(\varphi) \right|^2 \mathrm{d}\varphi \right)^{1/2}. \tag{2.38}$$

Proposition 2.1. *The value of* (2.38) *is invariant with respect to any rotation of the sensor location around the origin of the coordinate system.*

Proof. See Appendix A.1. □

Proposition 2.1 implies that the minimum of M_{MS} such that

$$\left\| a_n(\varphi) - S_{\frac{M_{\mathrm{MS}}-1}{2}}\{a_n\}(\varphi) \right\|_{L^2} \leq \epsilon \tag{2.39}$$

depends only on ϵ and the normalized sensor distance from the origin of the coordinate system, that is,

$$\frac{r_n}{\lambda} = \frac{\sqrt{p_{x_n}^2 + p_{y_n}^2}}{\lambda}. \tag{2.40}$$

In Figure 2.2, we depict the minimum of M_{MS} subject to the constraint (2.39). This figure shows that the required value of M_{MS} for the accurate approximation of the nth steering vector component is nearly an affine function of r_n/λ. Consequently, the required value of M_{MS} for the accurate approximation of the manifold is nearly an affine function of the largest sensor distance from the origin of the coordinate system [DD94a]. Note that in (2.4) we assumed that the phase center of the array coincides with the origin of the coordinate system. More generally, the required value of M_{MS} is nearly an affine function of the largest sensor distance from the phase center of the array.

2.2 Performance analysis of the array interpolation and MS techniques

In this section, we present an asymptotic performance analysis for the array interpolation and MS techniques [RG08a]. Our performance analysis is different from the earlier analysis presented in [BRK07] in that it takes into account both the finite sample effect and the array manifold approximation errors.[1] Our results are asymptotic, because they are based on the assumptions of high signal-to-noise ratio (SNR), large number of snapshots, and sufficiently small manifold approximation errors.

Proposition 2.2. *Suppose the signal and noise waveforms are white standard Gaussian random processes and the signal-subspace eigenvalues are pairwise distinct from each other. Then, the first-order approximation of the MSE of the DOA estimates from the MS technique is given by*

$$\begin{aligned}\mathrm{E}\{(\hat{\varphi}_l-\varphi_l)^2\} \approx {} & \frac{1}{\left(2\boldsymbol{d}^H_{\mathrm{MS},l}\boldsymbol{V}_{\boldsymbol{x},N}\boldsymbol{V}^H_{\boldsymbol{x},N}\boldsymbol{d}_{\mathrm{MS},l}\right)^2}\Bigg(4\Re^2\left\{\boldsymbol{d}^H_{\mathrm{MS},l}\boldsymbol{V}_{\boldsymbol{x},N}\boldsymbol{V}^H_{\boldsymbol{x},N}\boldsymbol{a}_{\mathrm{MS},l}\right\} \\ & + 8\,\Re\left\{\boldsymbol{d}^H_{\mathrm{MS},l}\boldsymbol{V}_{\boldsymbol{x},N}\boldsymbol{V}^H_{\boldsymbol{x},N}\boldsymbol{a}_{\mathrm{MS},l}\right\}\,\Re\left\{\sum_{k=1}^{L}\sum_{n=L+1}^{N}\mu_{k,n}\boldsymbol{d}^H_{\mathrm{MS},l}\left(\boldsymbol{v}_{\boldsymbol{x},k}\boldsymbol{v}^H_{\boldsymbol{x},k}-\boldsymbol{v}_{\boldsymbol{x},n}\boldsymbol{v}^H_{\boldsymbol{x},n}\right)\boldsymbol{a}_{\mathrm{MS},l}\right\} \\ & + 2\sum_{k=1}^{L}\sum_{n=L+1}^{N}\mu_{k,n}\left|\boldsymbol{d}^H_{\mathrm{MS},l}\boldsymbol{v}_{\boldsymbol{x},k}\boldsymbol{v}^H_{\boldsymbol{x},n}\boldsymbol{a}_{\mathrm{MS},l}+\boldsymbol{a}^H_{\mathrm{MS},l}\boldsymbol{v}_{\boldsymbol{x},k}\boldsymbol{v}^H_{\boldsymbol{x},n}\boldsymbol{d}_{\mathrm{MS},l}\right|^2\Bigg),\end{aligned} \tag{2.41}$$

where

$$\mu_{k,n}=\frac{\lambda_{\boldsymbol{x},k}\lambda_{\boldsymbol{x},n}}{K(\lambda_{\boldsymbol{x},k}-\lambda_{\boldsymbol{x},n})^2}, \tag{2.42}$$

and

$$\begin{aligned}\boldsymbol{a}_{\mathrm{MS},l} &= \boldsymbol{G}_{\mathrm{MS}}\,\boldsymbol{g}_{\mathrm{MS}}(\varphi_l) \\ \boldsymbol{d}_{\mathrm{MS},l} &= \boldsymbol{G}_{\mathrm{MS}}\,\frac{\partial\boldsymbol{g}_{\mathrm{MS}}(\varphi_l)}{\partial\varphi}.\end{aligned} \tag{2.43}$$

Proof. See Appendix A.2. ☐

Interestingly, similar steps as in Appendix A.2 show that replacing $\boldsymbol{a}_{\mathrm{MS},l}$ and $\boldsymbol{d}_{\mathrm{MS},l}$ with

$$\begin{aligned}\boldsymbol{a}_{\mathrm{AI},l} &= \boldsymbol{G}_{\mathrm{AI}}\,\boldsymbol{g}_{\mathrm{AI}}(\varphi_l) \\ \boldsymbol{d}_{\mathrm{AI},l} &= \boldsymbol{G}_{\mathrm{AI}}\,\frac{\partial\boldsymbol{g}_{\mathrm{AI}}(\varphi_l)}{\partial\varphi}\end{aligned} \tag{2.44}$$

---

[1] The analysis in [BRK07] does not take into account the finite sample effect, but instead it focusses on manifold approximation and array calibration errors.

in (2.41) gives the MSE performance of the array interpolation technique.

Note that if the manifold approximation errors in (2.23) and (2.28) are zero, then

$$\begin{aligned} \boldsymbol{a}_{\mathrm{MS},l} &= \boldsymbol{a}_{\mathrm{AI},l} = \boldsymbol{a}(\varphi_l) \\ \boldsymbol{d}_{\mathrm{MS},l} &= \boldsymbol{d}_{\mathrm{AI},l} = \frac{\partial \boldsymbol{a}(\varphi_l)}{\partial \varphi}. \end{aligned} \tag{2.45}$$

In this case, (2.41) can be simplified to

$$\mathrm{E}\left\{(\hat{\varphi}_l - \varphi_l)^2\right\} \approx \frac{\sigma}{2K} \frac{\sum_{k=1}^{L} \frac{\lambda_{\boldsymbol{x},k}}{(\lambda_{\boldsymbol{x},k}-\sigma)^2} \left|\boldsymbol{a}^H(\varphi_l)\boldsymbol{v}_{\boldsymbol{x},k}\right|}{\frac{\partial \boldsymbol{a}^H(\varphi_l)}{\partial \varphi} \boldsymbol{V}_{\boldsymbol{x},N} \boldsymbol{V}_{\boldsymbol{x},N}^H \frac{\partial \boldsymbol{a}(\varphi_l)}{\partial \varphi}}. \tag{2.46}$$

Therefore, the array interpolation and MS techniques lead to similar DOA estimation MSE performance if the errors in (2.23) and (2.28) are negligible as compared to the finite sample effect. Moreover, (2.46) is equivalent to the expression derived in [RH89] for the MSE performance of the conventional root-MUSIC estimator in the ULA case. That means, in the case of a ULA the conventional root-MUSIC and MS techniques lead to the same asymptotic performance if the approximation error in (2.28) is small. Hence, the root-MUSIC, interpolated root-MUSIC, and MS techniques asymptotically achieve similar DOA estimation performance, even though these methods estimate the DOAs by means of rooting different polynomials.

2.3 Fourier-domain root-MUSIC

In this section, we develop an alternative root-MUSIC-type approach to DOA estimation with arbitrary NUAs. The proposed approach is based on the observation that any polynomial approximation of the MUSIC null-spectrum function enables search-free DOA estimation. The polynomial coefficients should minimize the approximation error for the MUSIC null-spectrum function. For computational complexity reasons, the degree of the approximating polynomial should be as small as possible.

The Fourier series expansion of the MUSIC null-spectrum function can be expressed as

$$f_{\mathrm{MU}}(\varphi) = \sum_{m=-\infty}^{\infty} F_{f_{\mathrm{MU}}}[m]\, e^{jm\varphi}, \tag{2.47}$$

where the Fourier series coefficients are given by

$$F_{f_{\mathrm{MU}}}[m] = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{\mathrm{MU}}(\varphi)\, e^{-jm\varphi}\, \mathrm{d}\varphi. \tag{2.48}$$

Truncating the Fourier series in (2.47) to $2M_{\mathrm{FD}}-1$ points yields

$$f_{\mathrm{MU}}(\varphi) \approx \sum_{m=-(M_{\mathrm{FD}}-1)}^{M_{\mathrm{FD}}-1} F_{f_{\mathrm{MU}}}[m]\, e^{jm\varphi} = S_{M_{\mathrm{FD}}-1}\{f_{\mathrm{MU}}\}(\varphi). \tag{2.49}$$

The functions $S_{M_{\mathrm{FD}}-1}\{f_{\mathrm{MU}}\}(\varphi)$ and $f_{\mathrm{MS}}(\varphi)$ have the same structure if $M_{\mathrm{FD}} = M_{\mathrm{MS}}$. Then, these functions differ only with respect to the coefficients of the Fourier basis functions. The coefficient $F_{f_{\mathrm{MU}}}[m]$ of $S_{M_{\mathrm{FD}}-1}\{f_{\mathrm{MU}}\}(\varphi)$ corresponds to the coefficient

$$\sum_{\substack{k,l \\ l=k+m}} \left[\boldsymbol{G}_{\mathrm{MS}}^{H} \hat{\boldsymbol{V}}_{\boldsymbol{x},N} \hat{\boldsymbol{V}}_{\boldsymbol{x},N}^{H} \boldsymbol{G}_{\mathrm{MS}}\right]_{k,l} \tag{2.50}$$

of $f_{\mathrm{MS}}(\varphi)$.

Proposition 2.3. *The Fourier series coefficients $F_{f_{\mathrm{MU}}}[m]$ minimize the L^2-norm of the approximation error, i.e.,*

$$\left\| f_{\mathrm{MU}}(\varphi) - \sum_{m=-(M_{\mathrm{FD}}-1)}^{M_{\mathrm{FD}}-1} F_{f_{\mathrm{MU}}}[m] e^{jm\varphi} \right\|_{L^2} \leq \left\| f_{\mathrm{MU}}(\varphi) - \sum_{m=-(M_{\mathrm{FD}}-1)}^{M_{\mathrm{FD}}-1} H_m e^{jm\varphi} \right\|_{L^2} \tag{2.51}$$

for all H_m, and the equality holds if and only if

$$H_m = F_{f_{\mathrm{MU}}}[m], \quad m = -(M_{\mathrm{FD}}-1), \ldots, M_{\mathrm{FD}}-1. \tag{2.52}$$

Proof. See [Rud64], where this proposition is proven for a more general case. □

Proposition 2.3 implies that in terms of the L^2-norm $S_{M_{\mathrm{FD}}-1}\{f_{\mathrm{MU}}\}(\varphi)$ is closer or at least as close to the MUSIC null-spectrum function than $f_{\mathrm{MS}}(\varphi)$, provided that $M_{\mathrm{MS}} = M_{\mathrm{FD}}$.

The computational complexity to obtain the Fourier series coefficients may be too high for real-time applications. As an alternative, we have proposed in [RG08c] to use the DFT approximation of the Fourier series coefficients, that is,

$$\begin{aligned} &F_{f_{\mathrm{MU}}}[m] \approx \frac{1}{2\pi} \sum_{k=-(M_{\mathrm{FD}}-1)}^{M_{\mathrm{FD}}-1} f_{\mathrm{MU}}(k\Delta\varphi)\, e^{-jmk\,\Delta\varphi}\, \Delta\varphi \triangleq D_{f_{\mathrm{MU}}}[m], \\ &m = -(M_{\mathrm{FD}}-1), \ldots, M_{\mathrm{FD}}-1, \end{aligned} \tag{2.53}$$

where

$$\Delta\varphi = \frac{2\pi}{2M_{\mathrm{FD}}-1}. \tag{2.54}$$

Although in this case a close approximation of the Fourier series coefficients $F_{f_{\mathrm{MU}}}[m]$ can be achieved in a computationally efficient way, the resulting DFT coefficients $D_{f_{\mathrm{MU}}}[m]$ will be different from $F_{f_{\mathrm{MU}}}[m]$ due to aliasing effects introduced by sampling the MUSIC null-spectrum function. Using such a DFT approximation of $F_{f_{\mathrm{MU}}}[m]$, we obtain

$$S_{M_{\mathrm{FD}}-1}\{f_{\mathrm{MU}}\}(\varphi) \approx \sum_{m=-(M_{\mathrm{FD}}-1)}^{M_{\mathrm{FD}}-1} D_{f_{\mathrm{MU}}}[m] e^{jm\varphi} \triangleq f_{\mathrm{FD}}(\varphi). \tag{2.55}$$

The FD root-MUSIC polynomial is given by

$$f_{\mathrm{FD},\mathsf{z}}(\mathsf{z}) \triangleq \mathsf{z}^{M_{\mathrm{FD}}-1} \sum_{m=-(M_{\mathrm{FD}}-1)}^{M_{\mathrm{FD}}-1} D_{f_{\mathrm{MU}}}[m] \mathsf{z}^{m}, \tag{2.56}$$

where $\mathsf{z} = e^{j\varphi}$. The degree of the latter polynomial is $2M_{\mathrm{FD}} - 2$.

The roots of $f_{\mathrm{FD},\mathsf{z}}(\mathsf{z})$ satisfy the conjugate reciprocity property, i.e., if z_0 is a root of $f_{\mathrm{FD},\mathsf{z}}(\mathsf{z})$, then $\mathsf{z}_0' = 1/\mathsf{z}_0^*$ is also a root of this polynomial. To show this, let z_0 be a root of $f_{\mathrm{FD},\mathsf{z}}(\mathsf{z})$. Then,

$$\begin{aligned}
0 &= \mathsf{z}_0^{M_{\mathrm{FD}}-1} \sum_{m=-(M_{\mathrm{FD}}-1)}^{M_{\mathrm{FD}}-1} D_{f_{\mathrm{MU}}}[m] \mathsf{z}_0^{m} \\
&= \left(\mathsf{z}_0^{M_{\mathrm{FD}}-1}\right)^* \sum_{m=-(M_{\mathrm{FD}}-1)}^{M_{\mathrm{FD}}-1} \left(D_{f_{\mathrm{MU}}}[m]\right)^* \left(\mathsf{z}_0^{m}\right)^* \\
&= \left(\mathsf{z}_0^{M_{\mathrm{FD}}-1}\right)^* \sum_{m=-(M_{\mathrm{FD}}-1)}^{M_{\mathrm{FD}}-1} D_{f_{\mathrm{MU}}}[-m] \left(\mathsf{z}_0'\right)^{-m} \\
&= \left(\mathsf{z}_0^{M_{\mathrm{FD}}-1}\right)^* \sum_{m=-(M_{\mathrm{FD}}-1)}^{M_{\mathrm{FD}}-1} D_{f_{\mathrm{MU}}}[m] \left(\mathsf{z}_0'\right)^{m},
\end{aligned} \tag{2.57}$$

where we used for the third line that

$$\left(D_{f_{\mathrm{MU}}}[m]\right)^* = D_{f_{\mathrm{MU}}}[-m], \tag{2.58}$$

which directly follows from (2.53). Hence, $\mathsf{z}_0' = 1/\mathsf{z}_0^*$ is also a root of $f_{\mathrm{FD},\mathsf{z}}(\mathsf{z})$.

As the polynomials $f_{\mathrm{MS},\mathsf{z}}(\mathsf{z})$ and $f_{\mathrm{FD},\mathsf{z}}(\mathsf{z})$ have a similar structure, the required value of M_{FD} is approximately equal to that of M_{MS}.

The functions $f_{\mathrm{AI}}(\varphi)$ and $f_{\mathrm{MS}}(\varphi)$ are non-negative by definition, which is not the case for the function $f_{\mathrm{FD}}(\varphi)$. That is, $f_{\mathrm{FD}}(\varphi)$ may take values that are slightly below zero in some of its minima. Therefore, a different procedure has to be used to estimate

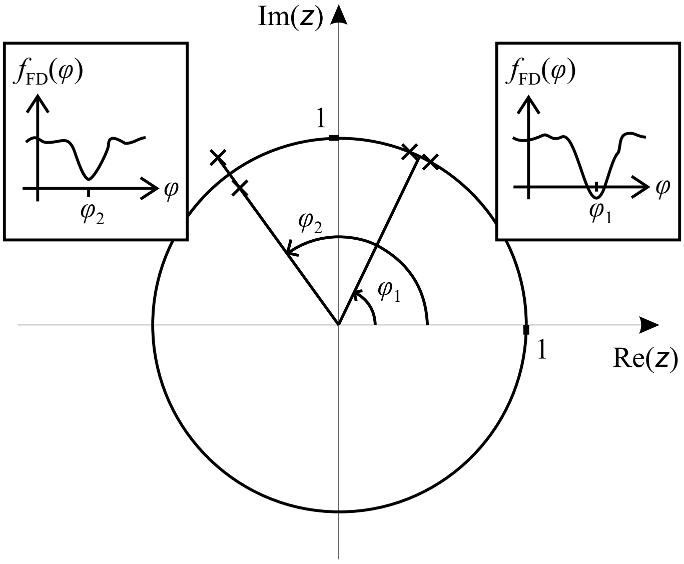


Figure 2.3. Signal root locations of the FD root-MUSIC polynomial in the case when $f_{\mathrm{FD}}(\varphi)$ has a negative minimum value (right part of the figure) or a positive minimum value (left part of the figure).

the signal DOAs from the roots of the polynomial $f_{\mathrm{FD,z}}(\mathsf{z})$ as compared to the array interpolation and MS techniques. The roots of $f_{\mathrm{FD,z}}(\mathsf{z})$ can be classified into two disjunct groups. The first group contains the roots that lie exactly on the unit circle. These roots are caused by the two sign changes of $f_{\mathrm{FD}}(\varphi)$ whenever $f_{\mathrm{FD}}(\varphi)$ takes a negative value. These roots appear in closely spaced pairs, but not in conjugate reciprocal pairs. The conjugate reciprocity property is trivially satisfied for these roots. The second group contains the roots that do not lie on the unit circle. These roots appear in conjugate reciprocal pairs and have the same nature as the roots of the conventional root-MUSIC polynomial.

In Figure 2.3, we schematically depict the root locations for these two different groups of roots. The root pairs of the first group of roots (i.e., roots located exactly on the unit circle) can be used to estimate the corresponding source DOAs by averaging the roots in each root pair. The remaining source DOAs can be estimated in the standard way (similar to that used in conventional root-MUSIC) from the roots of the second group.

Summarizing, the source DOAs can be estimated from the roots of $f_{\mathrm{FD,z}}(\mathsf{z})$ using the following procedure:

- **Step 1:** Take the root that is closest to the unit circle, and identify whether it belongs to the first or second group by checking if it is equal to its conjugate reciprocal value.

- **Step 2:** If the root belongs to the first group, then estimate the source DOA from the average of this root and its closest neighbor, and drop both these roots. Go to Step 4.
- **Step 3:** If the root belongs to the second group, then use it to estimate the source DOA, and drop both this root and its corresponding conjugate reciprocal root.
- **Step 4:** If less than L DOAs have been estimated, then go to Step 1. Otherwise, stop.

In [DD94b], it has been mentioned that the polynomial coefficients of the MS technique can be improved by choosing a large matrix $\boldsymbol{G}_{\mathrm{MS}}$, and then truncating the resulting polynomial. Similarly, the coefficients of the FD root-MUSIC polynomial can be improved by choosing a large DFT length, and then truncating the resulting polynomial. In both cases, the truncated polynomials may have negative values at the unit circle, and the proposed root selection algorithm should be used to estimate the source DOAs from the two groups of roots.

2.4 Weighted Fourier-domain root-MUSIC

As follows from Proposition 2.3, the Fourier series coefficients provide the best (in the L^2-norm sense) approximation of the MUSIC null-spectrum function $f_{\mathrm{MU}}(\varphi)$ within the class of complex trigonometric basis functions. However, to estimate the signal DOAs it is most important to have a close approximation of the MUSIC null-spectrum function in the vicinity of the true DOAs, i.e., in the areas where the null-spectrum has its minima. Using this idea, we take unequal positive weights for the samples of the function $f_{\mathrm{MU}}(\varphi)$, so that largest weights are assigned to smallest values of $f_{\mathrm{MU}}(\varphi)$ (that likely correspond to the angular areas where the sources are located) and, vice versa, lowest weights are assigned to largest values of $f_{\mathrm{MU}}(\varphi)$ [RG09a]. This leads to the following weighted least squares problem

$$\min_{\{H_m\}} \sum_{q=1}^{Q} w(\varphi_q) \left| f(\varphi_q) - \sum_{m=-(M_{\mathrm{WFD}}-1)}^{M_{\mathrm{WFD}}-1} H_m e^{jm\varphi_q} \right|^2, \tag{2.59}$$

where $w(\varphi_q)$ are the weight coefficients, and Q is the number of the MUSIC null-spectrum samples. To avoid that (2.59) is under-defined, we require that Q is greater than or equal to $2M_{\mathrm{WFD}}-1$. A natural choice for the weight coefficients to stress low values of $f_{\mathrm{MU}}(\varphi)$ is

$$w(\varphi_q)=\frac{1}{f(\varphi_q)},\quad q=1,\ldots,Q. \tag{2.60}$$

The optimization problem (2.59) can be written equivalently as

$$\min_{\boldsymbol{h}}\left\|\boldsymbol{W}_{\mathrm{WFD}}^{1/2}\left(\boldsymbol{f}-\boldsymbol{B}\,\boldsymbol{h}\right)\right\|_2^2, \tag{2.61}$$

where

$$\boldsymbol{W}_{\mathrm{WFD}}=\mathrm{diag}\{w(\varphi_1),\ldots,w(\varphi_Q)\} \tag{2.62}$$

is a diagonal matrix with diagonal entries $w(\varphi_1),\ldots,w(\varphi_Q)$,

$$\boldsymbol{f}=\left[f(\varphi_1),\ldots,f(\varphi_Q)\right]^T \tag{2.63}$$

contains the MUSIC null-spectrum samples,

$$\boldsymbol{h}=\left[H_{-M_{\mathrm{WFD}}+1},\ldots,H_{M_{\mathrm{WFD}}-1}\right]^T \tag{2.64}$$

contains the optimization variables,

$$\boldsymbol{B}=\begin{bmatrix} e^{-j(M_{\mathrm{WFD}}-1)\varphi_1} & e^{-j(M_{\mathrm{WFD}}-2)\varphi_1} & \ldots & e^{j(M_{\mathrm{WFD}}-1)\varphi_1}\\ e^{-j(M_{\mathrm{WFD}}-1)\varphi_2} & e^{-j(M_{\mathrm{WFD}}-2)\varphi_2} & \ldots & e^{j(M_{\mathrm{WFD}}-1)\varphi_2}\\ \vdots & \vdots & & \vdots\\ e^{-j(M_{\mathrm{WFD}}-1)\varphi_Q} & e^{-j(M_{\mathrm{WFD}}-2)\varphi_Q} & \ldots & e^{j(M_{\mathrm{WFD}}-1)\varphi_Q} \end{bmatrix}, \tag{2.65}$$

and $\|\cdot\|_2$ denotes the Euclidean vector norm. The solution of (2.61) is given by

$$\boldsymbol{h}^\star=\left(\boldsymbol{B}^H\,\boldsymbol{W}_{\mathrm{WFD}}\,\boldsymbol{B}\right)^{-1}\boldsymbol{B}^H\,\boldsymbol{W}_{\mathrm{WFD}}\,\boldsymbol{f}. \tag{2.66}$$

Let the optimum parameters of (2.59) be denoted by $H^\star_{-M_{\mathrm{WFD}}+1},\ldots,H^\star_{M_{\mathrm{WFD}}-1}$. The WFD approximation of the MUSIC null-spectrum function is

$$f_{\mathrm{WFD}}(\varphi)\triangleq\sum_{m=-(M_{\mathrm{WFD}}-1)}^{M_{\mathrm{WFD}}-1}H^\star_m e^{jm\varphi}. \tag{2.67}$$

Furthermore, the WFD root-MUSIC polynomial is defined as

$$f_{\mathrm{WFD},\mathsf{z}}(\mathsf{z})\triangleq\mathsf{z}^{M_{\mathrm{WFD}}-1}\sum_{m=-(M_{\mathrm{WFD}}-1)}^{M_{\mathrm{WFD}}-1}H^\star_m\mathsf{z}^m. \tag{2.68}$$

The degree of the latter polynomial is $2(M_{\mathrm{WFD}}-1)$. It can be verified easily that

$$H^\star_{-m}=\left(H^\star_m\right)^*. \tag{2.69}$$

Using this relation, it can be proved similar as in (2.57) that the roots of $f_{\mathrm{WFD},\mathsf{z}}(\mathsf{z})$ satisfy the conjugate reciprocity property. As the function $f_{\mathrm{WFD}}(\varphi)$ can take negative values, the same procedure as in FD root-MUSIC has to be used to estimate the signal DOAs from the roots of $f_{\mathrm{WFD},\mathsf{z}}(\mathsf{z})$.

Algorithm	Computational complexity
Spectral MUSIC	$\mathcal{O}(N^3 + J_{\mathrm{MU}}N^2)$
Interpolated root-MUSIC [Fri93]	$\mathcal{O}(N^3 + I_{\mathrm{AI}}(M_{\mathrm{AI}}N^2 + M_{\mathrm{AI}}^2 N + M_{\mathrm{AI}}^3))$
MS technique [DD94b, BRK07]	$\mathcal{O}(N^3 + MN^2 + M^2N + M^3) = \mathcal{O}(N^3 + M^3)$
FD root-MUSIC	$\mathcal{O}(N^3 + MN^2 + M\log M + M^3) = \mathcal{O}(N^3 + M^3)$
WFD root-MUSIC	$\mathcal{O}(N^3 + M^2Q + N^2Q + M^3)$

Table 2.1. Computational complexities of spectral MUSIC, interpolated root-MUSIC, the MS technique, and the FD and WFD root-MUSIC techniques.

2.5 Computational complexity

The asymptotic growth rates of the computational complexities of the proposed FD and WFD root-MUSIC algorithms are compared in Table 2.1 with those of the conventional spectral MUSIC, interpolated root-MUSIC, and MS techniques. The number of angular sectors in the interpolated root-MUSIC technique is denoted by I_{AI}. The degrees of the MS, FD root-MUSIC, and WFD root-MUSIC polynomials are assumed to be the same, i.e., $M_{\mathrm{MS}} = M_{\mathrm{FD}} = M_{\mathrm{WFD}} = M$.

All methods in Table 1 include the eigendecomposition step, which is represented by the term $\mathcal{O}(N^3)$ [GvL96]. Note that subspace tracking algorithms allow to recursively update the signal subspace eigenvectors with lower than cubic complexities [Ste92, Yan95, Str97]. The computation of J_{MU} samples of the MUSIC null-spectrum function requires $\mathcal{O}(J_{\mathrm{MU}}N^2)$ floating point operations. The complexities to compute the polynomial coefficients of the interpolated root-MUSIC, MS, FD root-MUSIC, and WFD root-MUSIC techniques are $\mathcal{O}(I_{\mathrm{AI}}(M_{\mathrm{AI}}N^2 + M_{\mathrm{AI}}^2 N))$, $\mathcal{O}(MN^2 + M^2N)$, $\mathcal{O}(MN^2 + M\log M)$, and $\mathcal{O}(M^2Q + N^2Q)$, respectively. The complexity to root a polynomial of degree M is $\mathcal{O}(M^3)$ [GvL96].

Table 1 shows that the computational complexity of FD root-MUSIC is comparable to that of the MS technique. It can be expected that $Q \in \Theta(M)$, so the complexity of WFD root-MUSIC is also comparable to that of the MS technique.

The results of Table 2.1 should be treated with care, because they specify only upper bounds on the asymptotic growth rates. Any constant factors are neglected, and it has not been taken into account that some of the parameters are closely related.

Let $r_{\max}$ denote the largest sensor distance from the origin of the coordinate system. Figure 2.2 shows that the required value of M is approximately proportional to $r_{\max}$. Hence, the complexities of the polynomial rooting steps of the MS, FD root-MUSIC, and WFD root-MUSIC techniques are $\mathcal{O}(r_{\max}^3)$. The growth of $r_{\max}$ with the number of sensors depends on the array geometry. For example, $r_{\max}$ is $\Theta(\sqrt{N})$ for URAs, $\Theta(N)$ for UCAs, and $\Theta(N^2)$ for linear MRAs [MD01].

Also, J_{MU} can be related to the number of sensors. Assuming a single source, it can be expected that J_{MU} is $\Theta(C_{\mathrm{CR}}^{-1/2})$, where C_{CR} is the minimum achievable MSE performance of any unbiased estimator of the source azimuth angle. If the source and noise waveforms are white standard Gaussian random processes, we have [vT02]

$$\begin{aligned} C_{\mathrm{CR}}^{-1} &= \frac{2P_1K}{\sigma}\frac{P_1N}{\sigma+P_1N}\frac{\partial \boldsymbol{a}^H(\varphi_1)}{\partial\varphi}\left(\boldsymbol{I}_N - \frac{\boldsymbol{a}(\varphi_1)\boldsymbol{a}^H(\varphi_1)}{\|\boldsymbol{a}(\varphi_1)\|_2^2}\right)\frac{\partial \boldsymbol{a}(\varphi_1)}{\partial\varphi} \\ &\le \frac{2P_1K}{\sigma}\left\|\frac{\partial \boldsymbol{a}(\varphi_1)}{\partial\varphi}\right\|_2^2 \\ &\le \frac{2P_1K}{\sigma}\left(\frac{2\pi}{\lambda}\right)^2 N r_{\max}^2, \end{aligned} \tag{2.70}$$

where P_1 denotes the source power, and the last line follows from (2.4). Consequently, if the number of snapshots and the SNR are bounded, then J_{MU} is $\mathcal{O}(\sqrt{N}r_{\max})$. The complexity of the spectral search step of the MUSIC algorithm is therefore $\mathcal{O}(N^{2.5}r_{\max})$.

The latter results suggest that the computational complexities of the MS, FD root-MUSIC, and WFD root-MUSIC techniques are higher than the computational complexity of the spectral MUSIC algorithm in the case of sparse arrays such as linear MRAs. In contrast, it can be expected that the computational complexities of the polynomial rooting steps of the MS, FD root-MUSIC, and WFD root-MUSIC techniques are lower than the complexity of the spectral search step of the conventional MUSIC algorithm in the case of dense planar arrays such as URAs.

2.6 Simulation results

In our first example, we validate the results of our asymptotic performance analysis of the MS technique. A randomly generated NUA of $N = 5$ sensors is used. The sensor locations are depicted in Figure 2.4. Two equally powered signals are assumed to impinge on the array from the directions $\varphi_1 = -10^\circ$ and $\varphi_2 = 20^\circ$. Throughout the following simulations, the source and noise waveforms are white standard Gaussian

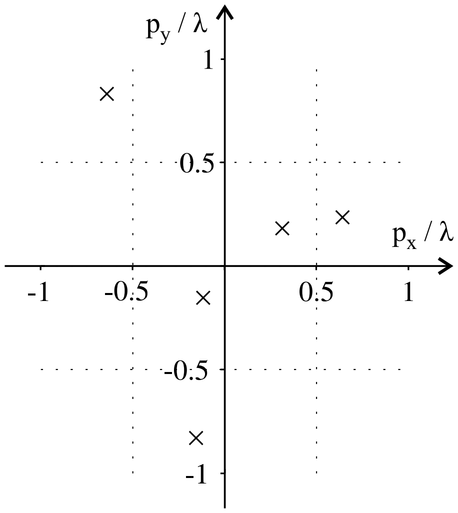


Figure 2.4. Array geometry.

random processes. The number of snapshots to estimate the array covariance matrix is $K = 100$.

Figure 2.5 displays the DOA estimation root mean squared errors (RMSEs) of the MS approach versus the SNR for different values of M_{MS}. The RMSE results are averaged over the sources. We compare the experimental performance obtained from 1000 independent Monte-Carlo runs with the analytical results of (2.41).

Figure 2.5 demonstrates that (2.41) accurately predicts the performance of the MS technique in the asymptotic domain. However, the first-order performance analysis is valid only for sufficiently small estimation errors (see Appendix A.2 for details). Hence, the results of (2.41) become inaccurate in the threshold domain (for SNR $<$ 0 dB). As can be observed from Figure 2.5, the performance curves saturate at high SNRs. This saturation is caused by the approximation error in (2.28). As can be expected, the negative effect of the approximation error reduces when increasing M_{MS}.

In our second example, we compare the DOA estimation performance of the spectral MUSIC, interpolated root-MUSIC, MS, FD root-MUSIC, and WFD root-MUSIC techniques. We assume two closely spaced sources that impinge on the array of Figure 2.4 from the directions $\varphi_1 = 15^\circ$ and $\varphi_2 = 20^\circ$. The spectral MUSIC technique searches over a uniform grid of $J_{\mathrm{MU}} = 1000$ null-spectrum samples. The interpolated root-MUSIC technique has been applied to sectors of width 60°. Virtual arrays with the same number of sensors as the real array have been selected, and their apertures have been chosen to be orthogonal to the center directions of these sectors. The MS, FD root-MUSIC, and WFD root-MUSIC techniques use $M_{\mathrm{MS}} = M_{\mathrm{FD}} = M_{\mathrm{WFD}} = M = 19$,

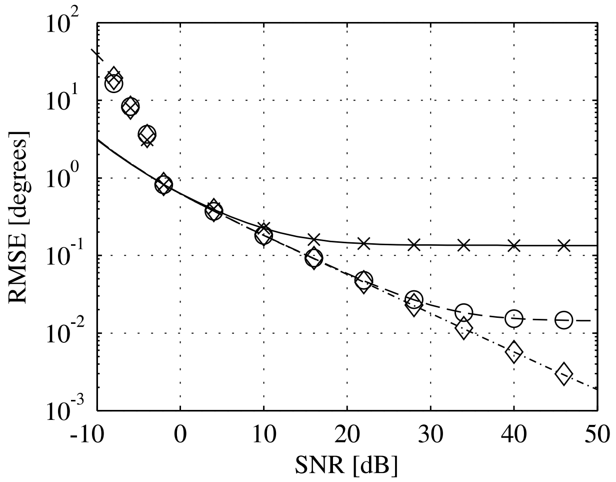


Figure 2.5. RMSE performance of the MS technique. Solid line: Analytical performance using $M_{\mathrm{MS}} = 19$. Crosses: Experimental performance using $M_{\mathrm{MS}} = 19$. Dashed line: Analytical performance using $M_{\mathrm{MS}} = 23$. Circles: Experimental performance using $M_{\mathrm{MS}} = 23$. Dash-dotted line: Analytical performance using $M_{\mathrm{MS}} = 27$. Diamonds: Experimental performance using $M_{\mathrm{MS}} = 27$.

i.e., the polynomials of these techniques have the same degree and differ only in their polynomial coefficients. The number of null-spectrum samples Q in the WFD root-MUSIC technique is twice the number of complex trigonometric functions used for the expansion. All other parameters are chosen as in the first example.

Figure 2.6 displays the DOA estimation RMSEs of the methods tested versus the SNR. The stochastic Cramér-Rao bound (CRB) is also shown [Kay93, vT02]. From this figure, it can be observed that the rooting based estimators have approximately the same threshold SNR, and that this threshold SNR is substantially better than that of the spectral MUSIC technique [RH89]. In the asymptotic domain, the FD and WFD root-MUSIC techniques outperform the MS technique. This suggests that the FD root-MUSIC and WFD root-MUSIC techniques achieve more accurate approximations of the MUSIC null-spectrum function than the MS technique in the vicinity of the source DOAs. Note that the asymptotic performance of all methods can be improved efficiently by means of a local minimization of the MUSIC null-spectrum function. Figure 2.6 demonstrates that the FD and WFD root-MUSIC techniques provide more accurate starting points for the local minimization than the MS technique.

In our third example, we examine the performance of the methods tested versus the angular separation between the sources. Figure 2.7 shows the DOA estimation RMSEs of the methods tested versus the source angular separation, where the DOA of the first source is varied while the DOA of the second source is fixed and equal to $\varphi_2 = 20°$.

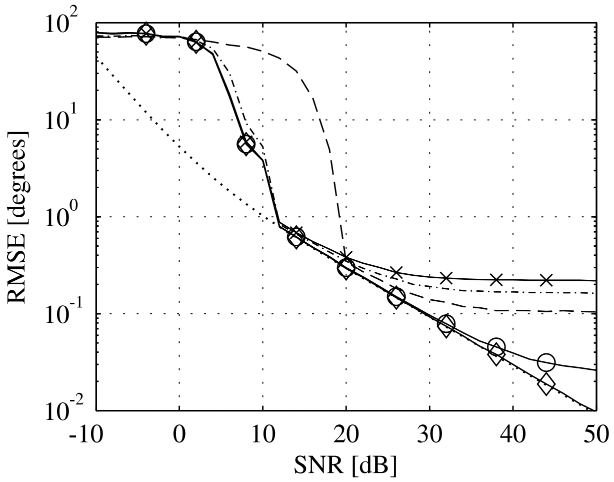


Figure 2.6. RMSE performance versus the SNR. Dashed line: Spectral MUSIC. Dash-dotted line: Interpolated root-MUSIC. Solid line with crosses: MS technique. Solid line with circles: FD root-MUSIC. Solid line with diamonds: WFD root-MUSIC. Dotted line: Stochastic CRB.

In this figure, SNR = 20 dB and all other parameters are chosen as in the previous example. Figure 2.7 demonstrates that FD and WFD root-MUSIC outperform the MS technique at medium/large source angular spacings.

Our fourth example studies the impact of the number of snapshots K on the methods tested. In this example, $\varphi_1 = 15°$ and $\varphi_2 = 20°$ are taken, and all other parameters are chosen as in the previous example. Figure 2.8 displays the DOA estimation RMSEs versus K. This figure shows that for high values of K we obtain rather similar results as for high SNR values, i.e., the proposed methods outperform the MS technique.

In our fifth example, we generalize the above results in the sense that we now consider the case when different (randomly generated) array geometries are used in each simulation run. In this example, the locations of $N = 6$ sensors have been drawn uniformly from the interior of a circle of radius λ. The performances of the FD root-MUSIC and the MS techniques are compared for SNR = 20 dB and $K = 100$ snapshots. The angular spacing between the two sources has been set to 20°. In Figure 2.9, a scatter plot for the DOA estimation RMSEs of the MS and FD root-MUSIC techniques is shown for $M = 15$. Each dot of this plot corresponds to one realization of the array geometry. Figure 2.9 demonstrates that FD root-MUSIC typically outperforms the MS technique. Note that the cases where FD root-MUSIC outperforms the MS technique correspond to all the dots that are located above the diagonal line.

In the last example, we study the impact of the parameter $M = M_{\mathrm{MS}} = M_{\mathrm{FD}} = M_{\mathrm{WFD}}$ on the performance of the methods tested. We use the same simulation settings as in

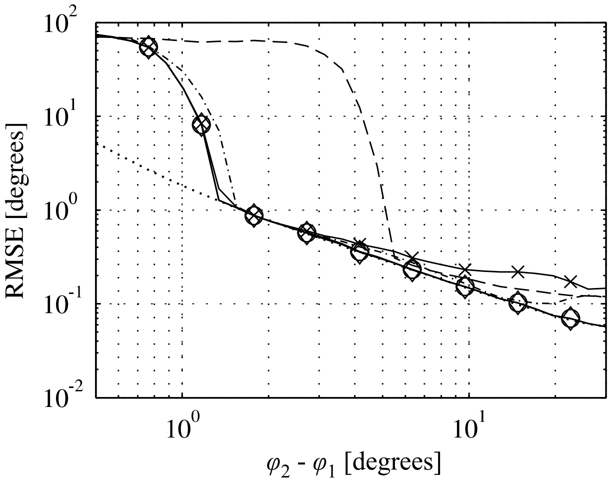


Figure 2.7. RMSE performance versus $\varphi_2 - \varphi_1$. Dashed line: Spectral MUSIC. Dash-dotted line: Interpolated root-MUSIC. Solid line with crosses: MS technique. Solid line with circles: FD root-MUSIC. Solid line with diamonds: WFD root-MUSIC. Dotted line: Stochastic CRB.

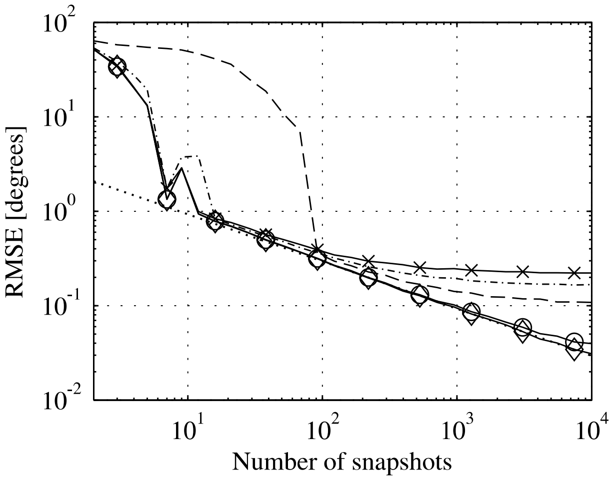


Figure 2.8. RMSE performance versus the number of snapshots. Dashed line: Spectral MUSIC. Dash-dotted line: Interpolated root-MUSIC. Solid line with crosses: MS technique. Solid line with circles: FD root-MUSIC. Solid line with diamonds: WFD root-MUSIC. Dotted line: Stochastic CRB.

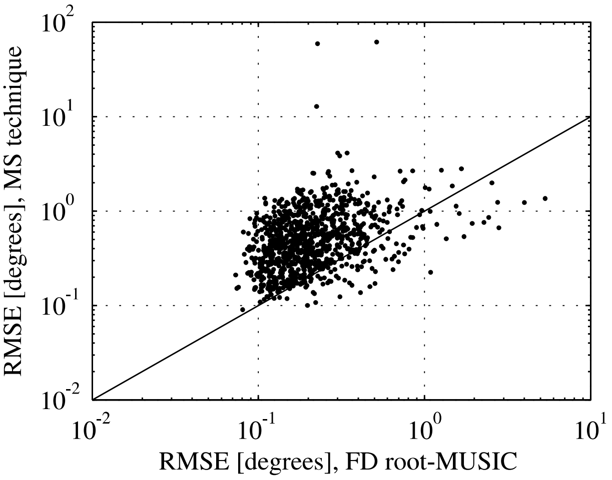


Figure 2.9. Scatter plot of the RMSEs of the MS and FD root-MUSIC techniques for randomly generated array geometries.

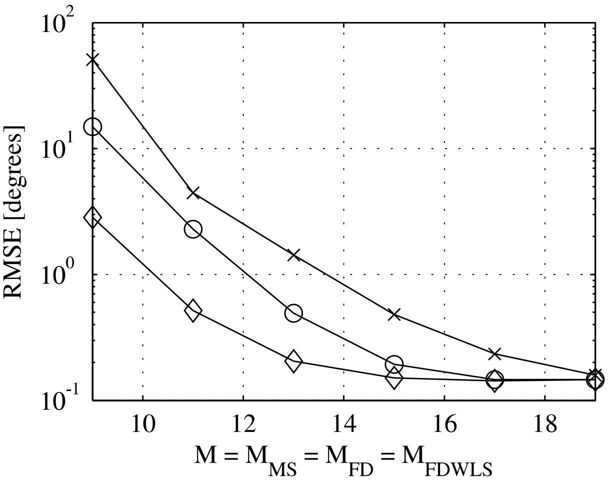


Figure 2.10. RMSE performance versus $M = M_{\mathrm{MS}} = M_{\mathrm{FD}} = M_{\mathrm{WFD}}$. Solid line with crosses: MS technique. Solid line with circles: FD root-MUSIC. Solid line with diamonds: WFD root-MUSIC. Dotted line: Stochastic CRB.

the previous example, and show the average RMSE performance over 1000 independent random realizations of the array geometry. Figure 2.10 clearly shows that FD root-MUSIC and WFD root-MUSIC require reduced degrees of their polynomials as compared to the MS technique to achieve the same performance.

In summary, the proposed DOA estimation algorithms are competitive alternatives to the MS technique in that these algorithms offer attractive performance-to-complexity tradeoffs.

2.7 Summary

In this chapter, the problem of spectral-search-free DOA estimation in arbitrary non-uniform sensor arrays has been addressed. First of all, an asymptotic performance analysis of the popular MS and interpolated root-MUSIC techniques has been presented.

Then, a root-MUSIC-type approach to DOA estimation in sensor arrays of arbitrary geometry has been developed. This approach has been referred to as Fourier-domain root-MUSIC as it is based on the DFT of the MUSIC null-spectrum samples.

To improve the DOA estimation performance of the proposed Fourier-domain root-MUSIC technique, a weighted least squares approximation of the MUSIC null-spectrum function has been developed.

It has been demonstrated through simulations with different array configurations that the proposed techniques offer attractive alternatives to the current state-of-the-art DOA estimation methods applicable to arrays with arbitrary geometries.

Chapter 3

Sparse array design for azimuthal DOA estimation

3.1 Background

A fundamental aspect of any array processing system is the array geometry as it strongly affects the system performance.

As in Chapter 2, we assume that only second-order statistics of the received data are used. In general, the spatial covariance function depends on the two locations at which the field is measured. However, in many applications the spatial field is wide-sense stationary. Then, the covariance function depends only on the displacement between the points at which the field is measured, but not on their absolute locations. Such wide-sense stationary spatial fields occur when the sources are uncorrelated and located in the far-field.

Uniform linear arrays of N sensors contain $N - m$ realizations of the displacement (lag) md_z, where d_z is the baseline. Hence, these arrays contain $N - m - 1$ redundancies for the lag md_z. With linear MRAs, the number of redundant lags is minimized [Mof68]. This leads to larger apertures as compared to ULAs with the same number of sensors and the same baseline d_z. Such increased apertures result in an improved spatial resolution capability at nearly the same hardware costs.

In the literature, so-called *general* and *restricted* linear MRAs have been studied [Mof68]. In the case of general linear MRAs, the geometry is optimized so that it contains the lags 0, d_z, $2d_z$ and so forth for as many contiguous integer multiples of the baseline d_z as possible. The aperture of general linear MRAs may be larger than the extent of the contiguous sequence of lags. In the case of restricted linear MRAs, the additional constraint that the set of lags does not contain any holes (missing integer multiples of d_z) is imposed. Hence, the aperture of restricted linear MRAs is equal to the extent of the contiguous sequence of lags.

Linear arrays of omnidirectional sensors provide only 180° angular coverage, and the spatial resolution capability is poor for directions close to the endfire direction. To achieve 360° angular coverage using omnidirectional sensors, planar array geometries are required. The MRA principle can be extended straightforwardly to planar geometries. However, linear and planar MRAs are particularly suitable to estimate the one-dimensional (1D) and two-dimensional (2D) source electric angles [TMS01]. Neither linear nor planar MRAs are optimized for the estimation of azimuth angles.

Since linear arrays provide good spatial resolution capability for limited angular sectors close to the broadside direction, arrays for azimuth estimation often consist of several linear arrays. These linear arrays are assembled to cover different angular sectors. Popular geometries are T-shaped, L-shaped, +-shaped, Y-shaped, and Δ-shaped arrays [Ell01, vT02]. Clearly, this approach is ad-hoc and does not satisfy an optimality criterion such as MRAs.

In this chapter, we propose a novel array geometry design for azimuthal DOA estimation [RG10b, RGb]. Our array geometries are appropriate for approximately planar systems such as terrestrial wireless communication systems or air traffic control systems. The proposed array design can be interpreted as a modification of the MRA concept for azimuth estimation.

Based on the proposed array geometry design, we develop an associated subspace-based method to estimate the DOAs of more sources than there are sensors (the case referred to in [AGGS98, ASG99a] as *"superior case DOA estimation"*). It is similar to the CA technique [PBNH85, AGGS98, ASG99a, ASG01, ASG03a, ASG03b], but in contrast to the latter technique, it provides non-ambiguous DOA estimates for the whole 360° azimuth field-of-view.

The proposed DOA estimation method first computes an augmented Toeplitz covariance matrix with corresponding Vandermonde steering vectors that contain in their elements the conventional Fourier basis functions. Subsequently, the source DOAs are estimated by applying a root-MUSIC-type algorithm to the Fourier-domain augmented covariance matrix. Hereafter, we refer to this technique as Fourier-domain covariance augmentation (FDCA) technique.

The proposed array geometry design leads to sparse planar arrays, for which manifold ambiguities can occur [MAK97, MP98, MPL99, ASG99b, MSD01]. Such manifold ambiguities can lead to a breakdown of subspace-based DOA estimation methods such as the MUSIC algorithm. We also demonstrate that the FDCA technique allows to resolve manifold ambiguities.

3.1.1 System model

Let us assume L narrowband far-field point sources that impinge on an array of N omnidirectional sensors. The baseband snapshot vector at the kth time instant is given by

$$\boldsymbol{x}(k) = \boldsymbol{A}(\boldsymbol{\varphi}, \boldsymbol{\theta})\boldsymbol{s}(k) + \boldsymbol{n}(k) \in \mathbb{C}^{N\times 1}, \tag{3.1}$$

where

$$\begin{aligned} \boldsymbol{\varphi} &= [\varphi_1, \ldots, \varphi_L]^T \in \mathbb{R}^{L\times 1} \\ \boldsymbol{\theta} &= [\theta_1, \ldots, \theta_L]^T \in \mathbb{R}^{L\times 1} \end{aligned} \tag{3.2}$$

are the vectors of azimuth and elevation angles, respectively,

$$\boldsymbol{A}(\boldsymbol{\varphi}, \boldsymbol{\theta}) = [\boldsymbol{a}(\varphi_1, \theta_1), \ldots, \boldsymbol{a}(\varphi_L, \theta_L)] \in \mathbb{C}^{N\times L} \tag{3.3}$$

is the steering matrix, $\boldsymbol{a}(\varphi_l, \theta_l)$ is the steering vector of the lth source,

$$\boldsymbol{s}(k) = [s_1(k), \ldots, s_L(k)]^T \in \mathbb{C}^{L\times 1} \tag{3.4}$$

is the source waveform vector, $s_l(k)$ is the complex envelope of the lth source, and $\boldsymbol{n}(k) \in \mathbb{C}^{N\times 1}$ is the noise vector. The steering vector of the lth source can be expressed as

$$\boldsymbol{a}(\varphi_l, \theta_l) = \begin{bmatrix} e^{j\frac{2}{\lambda}\boldsymbol{u}^T(\varphi_l,\theta_l)\boldsymbol{p}_1} \\ \vdots \\ e^{j\frac{2}{\lambda}\boldsymbol{u}^T(\varphi_l,\theta_l)\boldsymbol{p}_N} \end{bmatrix} \in \mathbb{C}^{N\times 1}, \tag{3.5}$$

where $\varphi_l \in [-\pi, \pi)$ and $\theta_l \in [-\pi/2, \pi/2]$ are the azimuth and elevation angles of the lth source, respectively,

$$\boldsymbol{u}(\varphi_l, \theta_l) = \begin{bmatrix} u_{x,l} \\ u_{y,l} \\ u_{z,l} \end{bmatrix} = \pi \begin{bmatrix} \cos\theta_l \cos\varphi_l \\ \cos\theta_l \sin\varphi_l \\ \sin\theta_l \end{bmatrix} \tag{3.6}$$

is a vector of length π, which points in the direction of the lth source,

$$\boldsymbol{p}_n = [p_{x_n}, p_{y_n}, p_{z_n}]^T \tag{3.7}$$

is the location of the nth sensor, and λ is the wavelength. The coordinate system is depicted in Figure 2.1. The components of $\boldsymbol{u}(\varphi_l, \theta_l)$ are referred to as the electric angles of the lth source. In the sequel, the array manifold and the number of sources are assumed to be known.

The baseband signal and noise waveforms are modeled as unknown zero-mean stationary uncorrelated random processes. Then, the $N \times N$ array covariance matrix can be expressed as

$$\boldsymbol{R}_{\boldsymbol{x}} = \mathrm{E}\{\boldsymbol{x}(k)\boldsymbol{x}^H(k)\} = \sum_{l=1}^{L} P_l \boldsymbol{a}(\varphi_l, \theta_l)\boldsymbol{a}^H(\varphi_l, \theta_l) + \sigma \boldsymbol{I}_N, \tag{3.8}$$

where $P_l = \mathrm{E}\{|s_l(k)|^2\}$ is the power of the lth signal, $\mathrm{E}\{\cdot\}$ denotes the statistical expectation, σ is the sensor noise power, and $\boldsymbol{I}_N$ is the $N \times N$ identity matrix. The (n_1, n_2)th component of $\boldsymbol{R}_{\boldsymbol{x}}$ can be expressed as

$$[\boldsymbol{R}_{\boldsymbol{x}}]_{n_1,n_2} = \sum_{l=1}^{L} P_l e^{j\frac{2}{\lambda}\boldsymbol{u}^T(\varphi_l,\theta_l)\left(\boldsymbol{p}_{n_1}-\boldsymbol{p}_{n_2}\right)} + \sigma\delta_{n_1,n_2}, \tag{3.9}$$

where δ_{n_1,n_2} is the Kronecker delta. Hence, the components of the array covariance matrix depend on the lags of the array. Let

$$\mathcal{S}_{\text{lag}} = \left\{\left(\boldsymbol{p}_{n_1} - \boldsymbol{p}_{n_2}\right) \mid n_1, n_2 \in \{1, \ldots, N\}\right\} \tag{3.10}$$

denote the set of all lags. Since there are $N-1$ redundancies for the zero-lag, the cardinality of $\mathcal{S}_{\text{lag}}$ is at most $N^2 - N + 1$. The geometry corresponding to $\mathcal{S}_{\text{lag}}$ is commonly referred to as the *co-array.*

In practice, the array covariance matrix can be estimated as

$$\hat{\boldsymbol{R}}_{\boldsymbol{x}} = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{x}(k)\boldsymbol{x}^H(k) = \boldsymbol{R}_{\boldsymbol{x}} + \boldsymbol{\Delta}_{\boldsymbol{R}_{\boldsymbol{x}}}, \tag{3.11}$$

where K is the number of snapshots, and $\boldsymbol{\Delta}_{\boldsymbol{R}_{\boldsymbol{x}}}$ is the covariance matrix estimation error.

3.1.2 Linear minimum redundancy arrays

Unless specified otherwise, we assume for linear arrays that the sensors lie on the z-axis of the coordinate system. In this case, we drop any dependencies on the azimuth angle for notational brevity.

The optimization problem for restricted linear MRAs can be formulated as

$$\begin{aligned} \max_{M,\boldsymbol{p}_z} \quad & M \\ \text{s.t.} \quad & \left\{ \tfrac{p_{z_{n_1}} - p_{z_{n_2}}}{d_z} \,\middle|\, n_1, n_2 \in \{1, \ldots, N\} \right\} = \{-M, \ldots, M\}, \end{aligned} \tag{3.12}$$

where

$$\boldsymbol{p}_z = [p_{z_1}, \ldots, p_{z_N}]^T$$

stacks the z-coordinates of the sensor locations. Hence, the constraint in (3.12) implies that the array geometry contains the lags $-Md_z, (-M+1)d_z, \ldots, Md_z$, and no other lags. Let M_{RL} denote the solution of (3.12), where the subscript refers to "restricted linear". In [Lee56] and [PPL90], it has been shown that

$$2.434 \leq \lim_{N \to \infty} \frac{N^2}{M_{\mathrm{RL}}} \leq 3. \tag{3.13}$$

Hence, M_{RL} increases quadratically with the number of sensors.

The optimum array geometry of (3.12) is not unique. It can be easily seen that the solution of (3.12) does not change if we introduce the constraints

$$\begin{aligned} &p_{z_1} < p_{z_2} < \ldots < p_{z_N}, \\ &p_{z_1} = 0, \quad p_{z_2} = d_z, \quad p_{z_N} = Md_z, \end{aligned} \tag{3.14}$$

where the constraints in the second line of (3.14) imply that the lag $(M-1)d_z$ is between the second and the Nth array sensors. The solution of (3.12) may be attained for several different geometries, even if the constraints in (3.14) are taken into account. Ambiguity-reducing constraints such as the ones in (3.14) are not essential for our approach, and therefore we do not specify them in the following.

The optimization problem for general linear MRAs can be formulated as

$$\begin{aligned} \max_{M, \boldsymbol{p}_z} \quad & M \\ \text{s.t.} \quad & \left\{ \tfrac{p_{z_{n_1}} - p_{z_{n_2}}}{d_z} \,\middle|\, n_1, n_2 \in \{1, \ldots, N\} \right\} \supseteq \{-M, \ldots, M\}, \end{aligned} \tag{3.15}$$

where $\mathcal{S}_1 \supseteq \mathcal{S}_2$ means that $\mathcal{S}_1$ is a superset of $\mathcal{S}_2$. The constraint in (3.15) implies that the array geometry contains the lags $-Md_z, (-M+1)d_z, \ldots, Md_z$, but it may contain also other lags. Let M_{GL} denote the solution of (3.15), where the subscript refers to "general linear". Since the constraint in (3.15) is less restrictive than the constraint in (3.12), we have for the same number of sensors

$$M_{\mathrm{GL}} \geq M_{\mathrm{RL}}. \tag{3.16}$$

It has been shown that [RR49, Lee56]

$$2.434 \leq \lim_{N \to \infty} \frac{N^2}{M_{\mathrm{GL}}} \leq 2.6646. \tag{3.17}$$

The complexity to compute linear MRA geometries increases exponentially with the number of sensors. Therefore, these geometries are known only for small numbers of sensors. In [LST93], restricted linear MRAs with $N \leq 17$ are listed. A list of general linear MRAs with $N \leq 11$ has been published in [Mof68].

The array covariance matrices of linear MRAs contain the entries

$$R_z(m) = \sum_{l=1}^{L} P_l e^{j\frac{2}{\lambda} m d_z u_{z,l}} + \sigma\delta_{m,0}, \quad m = -M, \dots, M, \tag{3.18}$$

where $M = M_{\mathrm{RL}}$ for restricted linear MRAs, and $M = M_{\mathrm{GL}}$ for general linear MRAs. In the following, we refer to linear MRAs with the baseline $d_z = \lambda/2$ as *standard* linear MRAs. For such arrays, (3.18) can be reformulated as

$$R_z(m) = \int_{-\pi}^{\pi} \Psi_{u_z}(u_z) e^{jmu_z}\, \mathrm{d}u_z + \sigma\delta_{m,0}, \quad m = -M, \dots, M, \tag{3.19}$$

where

$$\Psi_{u_z}(u_z) = \sum_{l=1}^{L} P_l \delta(u_z - u_{z,l}) \tag{3.20}$$

is the angular power density expressed as a function of u_z, and $\delta(\cdot)$ is the Dirac delta function. The formulation in (3.19) shows that the off-diagonal entries of the array covariance matrices of standard linear MRAs are Fourier series coefficients of $\Psi_{u_z}(u_z)$ scaled by 2π. This means that for standard linear MRAs, the number of available contiguous Fourier series coefficients of $\Psi_{u_z}(u_z)$ is maximized. Due to the orthogonality of the Fourier basis functions [Rud64], standard linear MRAs are particularly suitable to estimate the 1D source electric angles $u_{z,l}$.

3.1.3 Covariance augmentation technique

The array covariance matrices of linear MRAs contain the entries $R_z(-M), \dots, R_z(M)$. These arrays therefore enable to estimate the $(M+1) \times (M+1)$ augmented covariance matrix [PBNH85]

$$\boldsymbol{C}_{\mathrm{CA}} \triangleq \sum_{l=1}^{L} P_l \boldsymbol{a}_{\mathrm{CA}}(\theta_l) \boldsymbol{a}_{\mathrm{CA}}^H(\theta_l) + \sigma \boldsymbol{I}_M = \begin{bmatrix} R_z(0) & R_z(-1) & \cdots & R_z(-M) \\ R_z(1) & R_z(0) & & \vdots \\ \vdots & & \ddots & R_z(-1) \\ R_z(M) & \cdots & R_z(1) & R_z(0) \end{bmatrix}, \tag{3.21}$$

where

$$\boldsymbol{a}_{\mathrm{CA}}(\theta_l) \triangleq \left[1, e^{j\frac{2\pi}{\lambda} d_z \sin\theta_l}, \dots, e^{j\frac{2\pi}{\lambda} M d_z \sin\theta_l}\right]^T \tag{3.22}$$

is the $(M+1) \times 1$ augmented steering vector of the lth source. Applying the MUSIC algorithm (see Section 2.1.2) to the finite sample estimate of the augmented covariance matrix allows estimating the DOAs of up to M sources. Since $M > N$ if $N \geq 4$, the CA technique allows estimating the DOAs of more sources than there are sensors. For comparison, the conventional MUSIC algorithm allows estimating the DOAs of at most $N-1$ sources.

Generally, if the aperture of a linear array is larger than $(N-1)\lambda/2$, then manifold ambiguities exist [MP98]. That means, there exist DOAs $\theta_1, \ldots, \theta_L$ such that

$$\mathcal{R}(\boldsymbol{A}(\boldsymbol{\theta})) = \mathcal{R}(\boldsymbol{A}(\bar{\boldsymbol{\theta}})), \tag{3.23}$$

where

$$\bar{\boldsymbol{\theta}} = [\bar{\theta}_1, \ldots, \bar{\theta}_L]^T \tag{3.24}$$

is another vector of $L < N$ DOAs with at least one DOA that is not contained in $\boldsymbol{\theta}$, and $\mathcal{R}(\cdot)$ is the column-space operator. The ambiguity of the column-space of the steering matrix is critical for subspace-based DOA estimation methods. For example, the noise-free MUSIC null-spectrum function will have nulls at all angles θ_l and $\bar{\theta}_l$, $l = 1, \ldots, L$. The spurious nulls lead to outliers among the DOA estimates of the conventional MUSIC algorithm.

Since the apertures of linear MRAs are larger than $(N-1)\lambda/2$, manifold ambiguities exist for such arrays. However, the augmented steering vectors $\boldsymbol{a}_{\mathrm{CA}}(\theta_l)$ are non-ambiguous, because $L+1$ $(L+1 \leq N < M+1)$ different Vandermonde vectors of length $M+1$ are always linearly independent. Consequently,

$$\mathcal{R}(\boldsymbol{A}_{\mathrm{CA}}(\boldsymbol{\theta})) \neq \mathcal{R}(\boldsymbol{A}_{\mathrm{CA}}(\bar{\boldsymbol{\theta}})), \tag{3.25}$$

where

$$\boldsymbol{A}_{\mathrm{CA}}(\boldsymbol{\theta}) = [\boldsymbol{a}_{\mathrm{CA}}(\theta_1), \ldots, \boldsymbol{a}_{\mathrm{CA}}(\theta_L)] \tag{3.26}$$

is the $(M+1) \times L$ augmented steering matrix. The CA technique therefore allows to resolve manifold ambiguities [ASG99b].

3.1.4 Sparse planar arrays

Assuming omnidirectional sensors, planar arrays provide 360° azimuth coverage. In the case of planar arrays, we assume WLOG that the sensors are located in the xy-plane of the coordinate system. The baselines in x- and y-direction are denoted by d_x and d_y, respectively.

The MRA concept can be generalized straightforwardly to planar geometries. Similar to (3.12), the optimization problem for restricted planar MRAs can be formulated as

$$\begin{aligned} \max_{M,\boldsymbol{p}_{xy}} \quad & M \\ \text{s.t.} \quad & \left\{ \left(\tfrac{p_{x n_1} - p_{x n_2}}{d_x}, \tfrac{p_{y n_1} - p_{y n_2}}{d_y} \right) \,\middle|\, n_1, n_2 \in \{1, \dots, N\} \right\} = \{-M, \dots, M\}^2, \end{aligned} \tag{3.27}$$

where

$$\boldsymbol{p}_{xy} = [p_{x_1}, \dots, p_{x_N}, p_{y_1}, \dots, p_{y_N}]^T \tag{3.28}$$

contains the x- and y-coordinates of the sensor locations. Hence, the constraint in (3.27) enforces that the co-array has a uniform rectangular geometry with the baselines d_x and d_y in x- and y-directions, respectively. Let M_{RP} denote the solution of (3.27), where the subscript refers to “restricted planar”. It has been shown that [MD01]

$$2.207 \leq \lim_{N \to \infty} \frac{N}{M_{\mathrm{RP}}} \leq 3. \tag{3.29}$$

Hence, M_{RP} increases only linearly with the number of sensors.

The optimization problem for general planar MRAs is

$$\begin{aligned} \max_{M,\boldsymbol{p}_{xy}} \quad & M \\ \text{s.t.} \quad & \left\{ \left(\tfrac{p_{x n_1} - p_{x n_2}}{d_x}, \tfrac{p_{y n_1} - p_{y n_2}}{d_y} \right) \,\middle|\, n_1, n_2 \in \{1, \dots, N\} \right\} \supseteq \{-M, \dots, M\}^2. \end{aligned} \tag{3.30}$$

The solution M_{GP} (the subscript refers to “general planar”) of this problem is bounded asymptotically by [MD01]

$$2 \leq \lim_{N \to \infty} \frac{N}{M_{\mathrm{GP}}} \leq 2.6646, \tag{3.31}$$

where the lower bound follows from the fact that the cardinality of $\mathcal{S}_{\mathrm{lag}}$ is at most $N^2 - N + 1$.

The array covariance matrices of planar MRAs contain the spatial covariances

$$R_{xy}(m_1, m_2) = \sum_{l=1}^{L} P_l e^{j \frac{2}{\lambda}(m_1 d_x u_{x,l} + m_2 d_y u_{y,l})} + \sigma \delta_{m_1,0} \delta_{m_2,0}, \quad m_1, m_2 \in \{-M, \dots, M\}, \tag{3.32}$$

where $M = M_{\mathrm{RP}}$ for restricted planar MRAs, and $M = M_{\mathrm{GP}}$ for general planar MRAs. For standard (baselines $d_x = d_y = \lambda/2$) planar MRAs, the spatial covariances (3.32) can be written as

$$\begin{aligned} & R_{xy}(m_1, m_2) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \Psi_{u_x,u_y}(u_x, u_y) e^{j(m_1 u_x + m_2 u_y)} \, \mathrm{d}u_x \, \mathrm{d}u_y + \sigma \delta_{m_1,0} \delta_{m_2,0}, \\ & m_1, m_2 \in \{-M, \dots, M\}, \end{aligned} \tag{3.33}$$

where

$$\Psi_{u_x,u_y}(u_x, u_y) = \sum_{l=1}^{L} P_l \, \delta(u_x - u_{x,l}) \, \delta(u_y - u_{y,l}) \tag{3.34}$$

is the 2D angular power density expressed as a function of u_x and u_y. Hence, the off-diagonal entries of the array covariance matrices of standard planar MRAs are 2D Fourier series coefficients of $\Psi_{u_x,u_y}(u_x, u_y)$ scaled by $4\pi^2$. That means, for standard planar MRAs the number of available contiguous 2D Fourier series coefficients of $\Psi_{u_x,u_y}(u_x, u_y)$ is maximized. Standard planar MRAs are therefore particularly suitable to estimate the 2D source electric angles.

The complexity to compute planar MRA geometries is even higher than for linear MRAs with the same number of sensors [HRKV92]. Thus, sparse easy-to-compute geometries without holes among the set of lags have been proposed. The sparsity of such array can be assessed using (3.29).

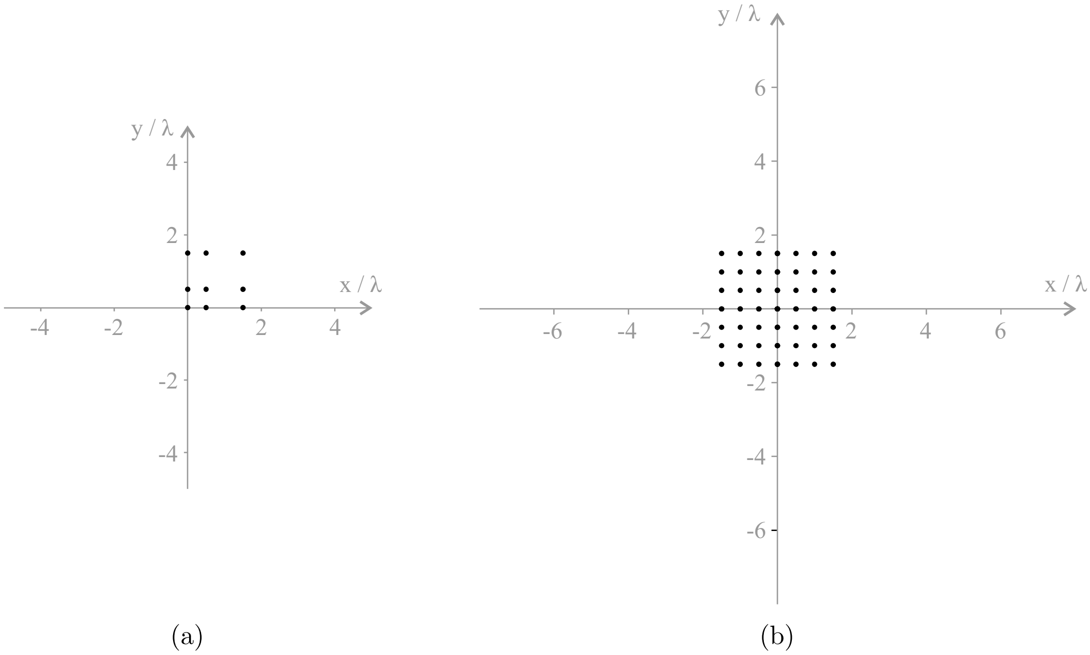
For example, cross-product arrays can be generated from two linear arrays, one in x-direction and the other one in y-direction, by placing a sensor at the location (x, y) if the first and second linear arrays contain sensors at $(x, 0)$ and $(0, y)$, respectively [Pum93]. The number of sensors of a cross-product array is equal to the product of the numbers of sensors of the two underlying linear arrays. Hence, cross-product arrays exist only for non-prime numbers of sensors. If restricted linear MRAs are used to generate a cross-product array, then the resulting co-array has a uniform rectangular geometry. Figure 3.1 depicts a cross-product array of $N = 9$ sensors.

Another class of sparse planar arrays has been proposed by Greene and Wood in [GW78]. These arrays have sensors at

$$\begin{aligned} &(md_x, 0), && m = 0, \ldots, M_{\mathrm{GW}} \\ &(0, md_y), && m = 1, \ldots, M_{\mathrm{GW}} \\ &(md_x, md_y), && m = 2, \ldots, M_{\mathrm{GW}}. \end{aligned} \tag{3.35}$$

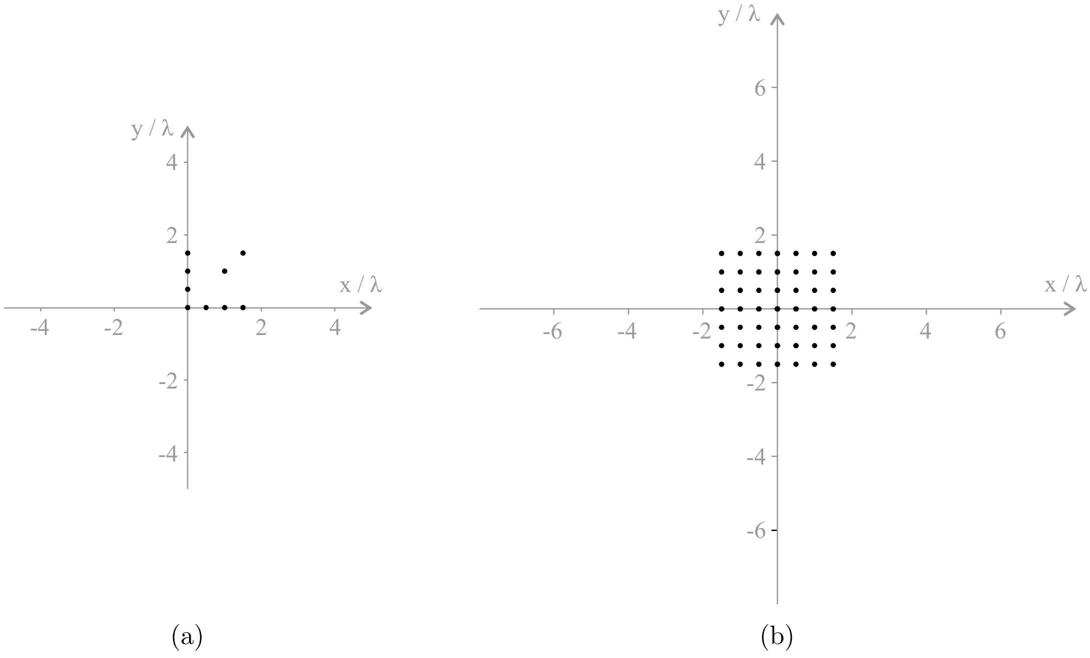
Hence, the number of sensors is $3M_{\mathrm{GW}}$. An example Greene-Wood array of $N = 9$ sensors is shown in Figure 3.2. The cross-product and Greene-Wood arrays of Figures 3.1(a) and 3.2(a) are restricted planar MRAs, but this is not generally true for cross-product and Greene-Wood arrays.

Intuitively, arrays for azimuth estimation should have an approximately circularly symmetric co-array. Uniform circular arrays are a popular class of planar arrays, which have this property [LP95]. However, UCAs sample the spatial field uniformly, and therefore they can be considered as the planar counterparts of ULAs. Figure 3.3 depicts a



(a) (b)

Figure 3.1. (a) Cross-product array of $N = 9$ sensors. (b) Corresponding co-array.



(a) (b)

Figure 3.2. (a) Greene-Wood array of $N = 9$ sensors. (b) Corresponding co-array.

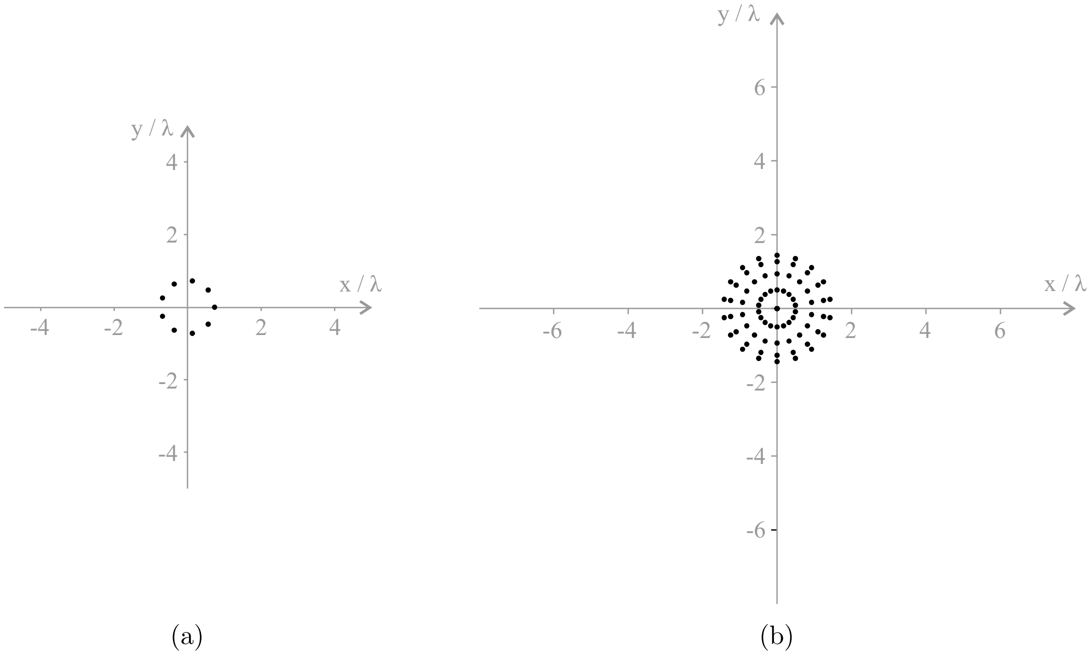


Figure 3.3. (a) Uniform circular array of $N = 9$ sensors. (b) Corresponding co-array.

UCA of $N = 9$ sensors and its corresponding co-array. The distance between adjacent sensors is chosen as $\lambda/2$ based on the phase-mode excitation theory [MZ94, vT02].

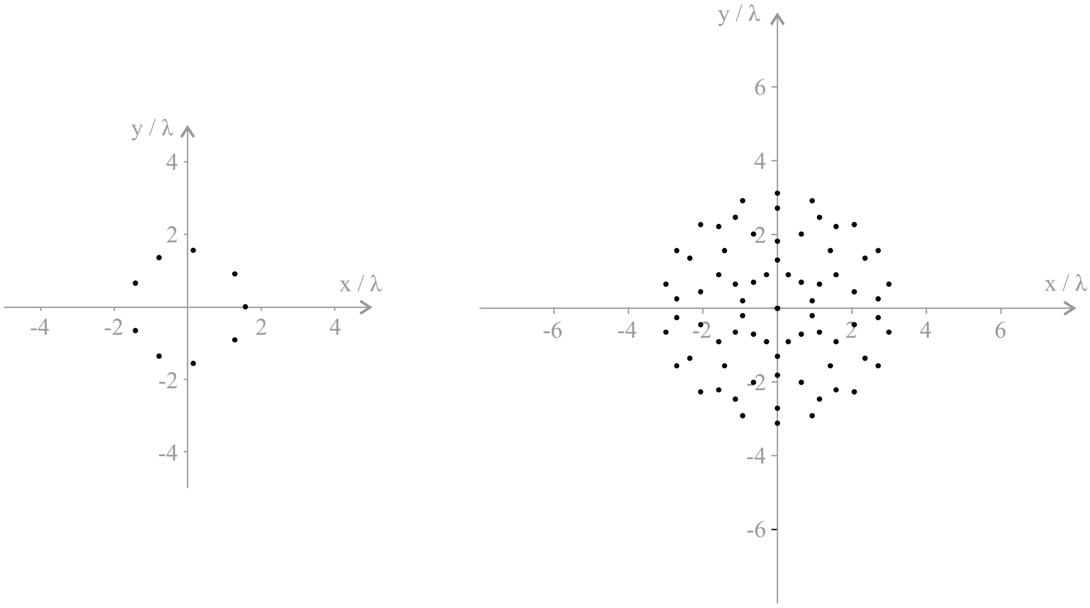
Another array geometry design, which leads to approximately circularly symmetric co-arrays, has been proposed by Cornwell in [Cor88]. The sensors are distributed inside a circle such that the product of the distances between the points of the co-array is maximized. This prevents small distances between the points of the co-array, and leads to a rather uniform sampling of the spatial covariance function. The array geometry optimization problem of [Cor88] can be formulated as

$$\begin{aligned} \max_{\boldsymbol{p}_{xy}} \quad & d_{\mathrm{Co}}(\boldsymbol{p}_{xy}) \\ \text{s.t.} \quad & p_{x_n}^2 + p_{y_n}^2 \leq 1, \quad n = 1, \ldots, N, \end{aligned} \tag{3.36}$$

where

$$d_{\mathrm{Co}}(\boldsymbol{p}_{xy}) = \sum_{\substack{n_1,n_2,n_3,n_4 \\ (n_1,n_2)\neq(n_3,n_4) \\ n_1\neq n_2}} \log \left\| \begin{bmatrix} (p_{x_{n_1}} - p_{x_{n_2}}) - (p_{x_{n_3}} - p_{x_{n_4}}) \\ (p_{y_{n_1}} - p_{y_{n_2}}) - (p_{y_{n_3}} - p_{y_{n_4}}) \end{bmatrix} \right\|_2 , \tag{3.37}$$

the cases $(n_1, n_2) \neq (n_3, n_4)$ and $n_1 \neq n_2$ are excluded from the summation to prevent zero-arguments of the logarithm, and $\|\cdot\|_2$ denotes the Euclidean vector norm. Empirical results show that the optimum sensor locations of (3.36) satisfy the constraints with equality, and that the co-arrays have a regular structure [Cor88]. Figure 3.4 depicts a



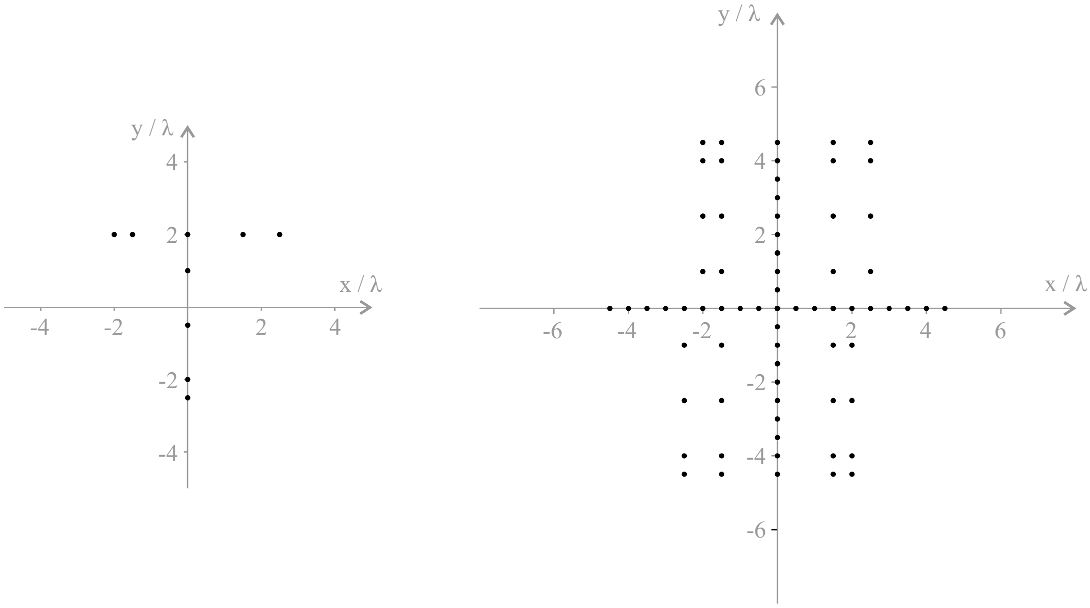
(a) (b)

Figure 3.4. (a) Cornwell array of $N = 9$ sensors. (b) Corresponding co-array.

Cornwell array of $N = 9$ sensors and the corresponding co-array. The sensor locations are scaled such that the average nearest neighbor distance between the points of the co-array is equal to $\lambda/2$.

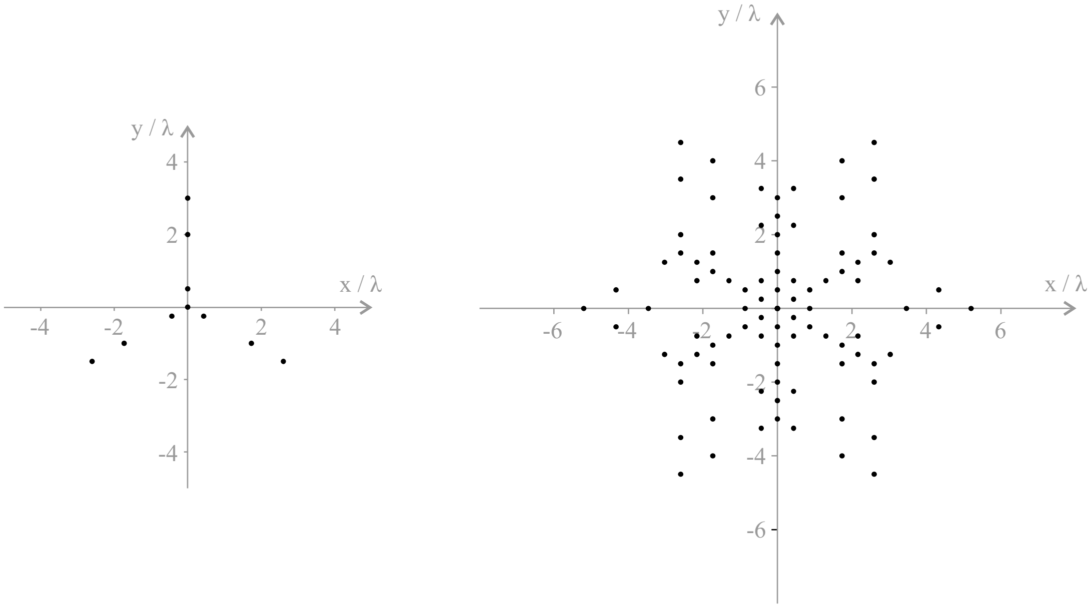
Planar MRAs as well as Cornwell arrays are designed to achieve a uniform sampling of the spatial covariance function. Since the cardinality of $\mathcal{S}_{\text{lag}}$ increases at most quadratically with the number of sensors, the apertures of planar MRAs and Cornwell arrays increase only linearly with the number of sensors. The uniform planar sampling of the spatial covariance function is well-justified for the estimation of 2D source electric angles, but we will see in the following that it is non-optimum for the estimation of azimuth angles.

Since linear arrays provide rather good estimation performance for limited angular sectors, practically used arrays for azimuth estimation often consist of several linear arrays. The linear arrays are assembled to cover different angular sectors. Popular geometries are T-shaped, L-shaped, +-shaped, Y-shaped, and Δ-shaped arrays [Ell01, vT02]. If these arrays are built using linear MRAs, then the array apertures increase quadratically with the number of sensors due to (3.13). Figures 3.5 and 3.6 show example T-shaped and Y-shaped arrays. However, this approach is obviously ad-hoc and does not satisfy an optimality criterion such as MRAs.



(a) (b)

Figure 3.5. (a) T-shaped array of $N = 9$ sensors, consisting of two standard restricted linear MRAs ·1·3·3·2·, where the numbers between the dots determine the separations between adjacent sensors in multiples of $\lambda/2$. (b) Corresponding co-array.



(a) (b)

Figure 3.6. (a) “Y”-shaped array of $N = 10$ sensors, consisting of three standard linear MRAs ·1·3·2·, where the numbers between the dots determine the separations between adjacent sensors in multiples of $\lambda/2$. (b) Corresponding co-array.

3.2 Proposed array geometry design

In Section 3.1.2, we showed that for standard linear MRAs, the number of available contiguous Fourier series coefficients of $\Psi_{u_z}(u_z)$ is maximized. Therefore, standard linear MRAs are particulary suitable to estimate the 1D source electric angles $u_{z,l}$. For a similar reason, standard planar MRAs are particularly suitable to estimate the 2D source electric angles $(u_{x,l}, u_{y,l})$. In this section, we propose a novel approach to the design of planar arrays for 360° azimuth estimation.

Towards this end, we maximize the number of available contiguous Fourier series coefficients of the azimuthal power density function

$$\Psi_\varphi(\varphi) = \sum_{l=1}^{L} P_l \delta(\varphi - \varphi_l). \tag{3.38}$$

Assuming that the sources and sensors are located in the xy-plane of the coordinate system, the (n_1, n_2)th component of $\boldsymbol{R_x}$ can be expressed as

$$[\boldsymbol{R_x}]_{n_1,n_2} = \sum_{l=1}^{L} P_l e^{j\frac{2\pi}{\lambda}\left(\left(p_{x_{n_1}} - p_{x_{n_2}}\right)\cos\varphi_l + \left(p_{y_{n_1}} - p_{y_{n_2}}\right)\sin\varphi_l\right)} + \sigma\delta_{n_1,n_2}. \tag{3.39}$$

Equation (3.39) can be reformulated as

$$[\boldsymbol{R_x}]_{n_1,n_2} = \left\langle e^{j\frac{2\pi}{\lambda}\left(\left(p_{x_{n_1}} - p_{x_{n_2}}\right)\cos\varphi + \left(p_{y_{n_1}} - p_{y_{n_2}}\right)\sin\varphi\right)}, \Psi_\varphi(\varphi)\right\rangle + \sigma\delta_{n_1,n_2}, \tag{3.40}$$

where

$$\langle f(\varphi), g(\varphi)\rangle \triangleq \int_{-\pi}^{\pi} f(\varphi) g^*(\varphi) \mathrm{d}\varphi \tag{3.41}$$

is the scalar product for the Hilbert space $L^2([-\pi, \pi])$ of Lebesgue measurable functions with finite norm [Rud64]

$$\|f(\varphi)\|_{L^2} \triangleq \left(\int_{-\pi}^{\pi} |f(\varphi)|^2 \mathrm{d}\varphi\right)^{1/2}. \tag{3.42}$$

The mth Fourier series coefficient of $\Psi_\varphi(\varphi)$ can be expressed as

$$F_{\Psi_\varphi}[m] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Psi_\varphi(\varphi) e^{-jm\varphi} \mathrm{d}\varphi = \left\langle \frac{1}{2\pi} e^{-jm\varphi}, \Psi_\varphi(\varphi)\right\rangle. \tag{3.43}$$

Clearly, it is not possible to choose the sensor locations such that (3.40) equals a Fourier series coefficient of $\Psi_\varphi(\varphi)$ or a scaled version thereof.

We use linear functions of $\hat{\boldsymbol{R}}_{\boldsymbol{x}}$ to estimate the Fourier series coefficients of $\Psi_\varphi(\varphi)$. To eliminate the disturbing effect of the non-vanishing (for $K \to \infty$) sensor noise component in the diagonal entries of $\hat{\boldsymbol{R}}_{\boldsymbol{x}}$, we define $\underline{\mathrm{vec}}\{\cdot\}$ as the vectorization operator for a square matrix, which leaves out the entries on the main diagonal. Then, the estimator for the mth Fourier series coefficient of $\Psi_\varphi(\varphi)$ can be expressed as

$$\hat{F}_{\Psi_\varphi}[m] = \boldsymbol{w}_m^H \underline{\mathrm{vec}}\{\hat{\boldsymbol{R}}_{\boldsymbol{x}}\}, \tag{3.44}$$

where the $N^2 - N \times 1$ weight vector $\boldsymbol{w}_m$ is yet to be determined. Using (3.11), we have

$$\hat{F}_{\Psi_\varphi}[m] = \boldsymbol{w}_m^H \underline{\mathrm{vec}}\{\boldsymbol{R}_{\boldsymbol{x}}\} + \boldsymbol{w}_m^H \underline{\mathrm{vec}}\{\boldsymbol{\Delta}_{\boldsymbol{R}_{\boldsymbol{x}}}\}. \tag{3.45}$$

We can write

$$\underline{\mathrm{vec}}\{\boldsymbol{R}_{\boldsymbol{x}}\} = \sum_{l=1}^{L} P_l \boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi_l) = \left\langle \boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi), \Psi_\varphi(\varphi) \right\rangle, \tag{3.46}$$

where

$$\boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi) = \underline{\mathrm{vec}}\{\boldsymbol{a}(\varphi)\boldsymbol{a}^H(\varphi)\} \tag{3.47}$$

is an $N^2 - N \times 1$ vector, the subscript $\boldsymbol{p}_{xy}$ emphasizes the dependency on the array geometry, and the scalar product of $\boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi)$ and $\Psi_\varphi(\varphi)$ is defined element-wise. Using (3.46), we obtain

$$\hat{F}_{\Psi_\varphi}[m] = \left\langle \boldsymbol{w}_m^H \boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi), \Psi_\varphi(\varphi) \right\rangle + \boldsymbol{w}_m^H \underline{\mathrm{vec}}\{\boldsymbol{\Delta}_{\boldsymbol{R}_{\boldsymbol{x}}}\}. \tag{3.48}$$

To estimate the mth Fourier series coefficient of $\Psi_\varphi(\varphi)$, we choose $\boldsymbol{w}_m$ such that

$$\boldsymbol{w}_m^H \boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi) = \frac{1}{2\pi} e^{-jm\varphi} + \delta_m(\varphi), \tag{3.49}$$

where $\delta_m(\varphi)$ is a small approximation error. Substituting (3.49) in (3.48) yields

$$\hat{F}_{\Psi_\varphi}[m] = F_{\Psi_\varphi}[m] + \langle \delta_m(\varphi), \Psi_\varphi(\varphi) \rangle + \boldsymbol{w}_m^H \underline{\mathrm{vec}}\{\boldsymbol{\Delta}_{\boldsymbol{R}_{\boldsymbol{x}}}\}. \tag{3.50}$$

Note that $\delta_m(\varphi)$ does not depend on the received data. Hence, it leads to biased estimates of the Fourier series coefficients of $\Psi_\varphi(\varphi)$. The variance of the estimation errors is solely due to the finite sample effect.

The error terms in (3.50) depend on $\Psi_\varphi(\varphi)$ and $\boldsymbol{\Delta}_{\boldsymbol{R}_{\boldsymbol{x}}}$. Since these quantities are unknown a-priori, we use the constraints

$$\left\| \boldsymbol{w}_m^H \boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi) - \frac{1}{2\pi} e^{-jm\varphi} \right\|_{L^2} \leq \epsilon, \quad m = 1, \ldots, M \tag{3.51}$$

and

$$\frac{1}{M} \sum_{m=1}^{M} \|\boldsymbol{w}_m\|_2 \leq \xi \tag{3.52}$$

to achieve small estimation errors for the Fourier series coefficients of $\Psi_\varphi(\varphi)$, where ϵ and ξ are user-defined parameters. For the choice of ξ, an analogy to standard linear MRAs can be exploited. For such arrays, the Fourier series coefficients of $\Psi_{u_z}(u_z)$ can be estimated as

$$\hat{F}_{\Psi_{u_z}}[m] = \tilde{\boldsymbol{w}}_m^H \underline{\mathrm{vec}}\{\hat{\boldsymbol{R}}_{\boldsymbol{x}}\} = F_{\Psi_{u_z}}[m] + \tilde{\boldsymbol{w}}_m^H \underline{\mathrm{vec}}\{\boldsymbol{\Delta}_{\boldsymbol{R}_{\boldsymbol{x}}}\}, \tag{3.53}$$

where $\tilde{\boldsymbol{w}}_m$ is an $N^2 - N \times 1$ selection vector scaled by $1/(2\pi)$. Hence, if $\xi = 1/(2\pi)$, then the finite sample errors in (3.50) can be expected to be comparable to those in (3.53).

Since $\Psi_\varphi(\varphi)$ is real-valued, its $(-m)$th Fourier series coefficient equals the complex conjugate of the mth Fourier series coefficient. Therefore, the constraints in (3.51) and (3.52) support estimation of the Fourier series coefficients for all $1 \leq |m| \leq M$.

To maximize the number of available contiguous Fourier series coefficients of $\Psi_\varphi(\varphi)$, we jointly optimize the weight vectors and sensor locations as

$$\begin{aligned} \max_{M, \boldsymbol{p}_{xy}, \{\boldsymbol{w}_m\}} \quad & M \\ \text{s.t.} \quad & \left\| \boldsymbol{w}_m^H \boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi) - \frac{1}{2\pi} e^{-jm\varphi} \right\|_{L^2} \leq \epsilon, \quad m = 1, \ldots, M \\ & \frac{1}{M} \sum_{m=1}^{M} \|\boldsymbol{w}_m\|_2 \leq \xi. \end{aligned} \tag{3.54}$$

To avoid the trivial solution, we assume in the sequel that $\epsilon < 1/\sqrt{2\pi}$.

Remark 3.1. Let M_{FD} (the subscript refers to "Fourier-domain") denote the solution of (3.54).[1] The components of $\boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi)$ span a sub-space of dimension at most $N^2 - N$. As the functions $\exp(-jm\varphi)$ are orthogonal, we have

$$2M_{\mathrm{FD}} \leq N^2 - N. \tag{3.55}$$

Remark 3.2. The more Fourier series coefficients of $\Psi_\varphi(\varphi)$ are available, the better is the angular resolution capability. As the angular resolution capability is closely related to the array aperture [vT02], maximizing M implicitly leads to large apertures.

Remark 3.3. Due to the constraints in (3.27) and (3.30), the array covariance matrices of standard planar MRAs contain the entries

$$\left\langle e^{j\pi(m_1 \cos\varphi + m_2 \sin\varphi)}, \Psi_\varphi(\varphi) \right\rangle, \quad m_1, m_2 \in \{-M, \ldots, M\}, \ (m_1, m_2) \neq (0,0). \tag{3.56}$$

The functions

$$e^{j\pi(m_1 \cos\varphi + m_2 \sin\varphi)}, \quad m_1, m_2 \in \{-M, \ldots, M\}, \ (m_1, m_2) \neq (0,0) \tag{3.57}$$

---

[1] This parameter should not be confused with the parameter M_{FD} used in Chapter 2.

can be accurately approximated by truncated Fourier series expansions (see Section 2.1.5). Using these orthogonal expansions, it can be easily verified that the functions in (3.57) are practically linearly dependent for $M \geq 3$. Therefore, even if a standard planar MRA does not contain any non-trivial redundancies, the off-diagonal entries of the array covariance matrix are still redundant for azimuth estimation. In other words, the constraints in (3.27) and (3.30) enforce that the spatial covariance function is estimated for a set of lags, which is redundant for azimuth estimation.

Remark 3.4. The proposed array geometry design (3.54) can be modified straightforwardly to the case of directional sensors. Then, also the sensor orientations have to be optimized.

3.2.1 Implementation

The problem in (3.54) is a mixed integer non-linear programming problem. We approach this problem by solving

$$\begin{aligned} \min_{\boldsymbol{p}_{xy}, \{\boldsymbol{w}_m\}} \quad & \frac{1}{M} \sum_{m=1}^{M} \|\boldsymbol{w}_m\|_2 \\ \text{s.t.} \quad & \left\| \boldsymbol{w}_m^H \boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi) - \frac{1}{2\pi} e^{-jm\varphi} \right\|_{L^2} \leq \epsilon, \quad m = 1, \dots, M \end{aligned} \tag{3.58}$$

for different values of M. Then, M_{FD} is the maximum value of M such that the solution of (3.58) is less than or equal to ξ.

The problem in (3.58) is convex with respect to the weight vectors, but non-convex with respect to the sensor locations. We divide this problem into an outer optimization of the sensor locations, and an inner optimization of the weight vectors. Let $f_M(\boldsymbol{p}_{xy})$ denote the solution of the inner optimization problem

$$\begin{aligned} \min_{\{\boldsymbol{w}_m\}} \quad & \frac{1}{M} \sum_{m=1}^{M} \|\boldsymbol{w}_m\|_2 \\ \text{s.t.} \quad & \left\| \boldsymbol{w}_m^H \boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi) - \frac{1}{2\pi} e^{-jm\varphi} \right\|_{L^2} \leq \epsilon, \quad m = 1, \dots, M. \end{aligned} \tag{3.59}$$

To determine the sensor locations that minimize $f_M(\boldsymbol{p}_{xy})$, we generate a large number of random array geometries $\boldsymbol{p}_{xy}$. For each of these arrays, we first determine the optimum scaling, and subsequently perform a local optimization of the best scaled arrays.

For the inner optimization problem, the sensor locations are fixed. To evaluate the L^2-norms in (3.58), we sample the functions at

$$\varphi_k = 2\pi k / N_\varphi, \quad k = 0, \dots, N_\varphi - 1. \tag{3.60}$$

Using the notations

$$\boldsymbol{g}_m = \frac{1}{2\pi} \begin{bmatrix} e^{jm\varphi_0} & \dots & e^{jm\varphi_{N_\varphi - 1}} \end{bmatrix}^T \in \mathbb{C}^{N_\varphi \times 1} \tag{3.61}$$

and

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi_0) & \dots & \boldsymbol{b}_{\boldsymbol{p}_{xy}}\left(\varphi_{N_\varphi - 1}\right) \end{bmatrix} \in \mathbb{C}^{N^2 - N \times N_\varphi}, \tag{3.62}$$

the constraints in (3.58) can be formulated as

$$\left\| \boldsymbol{w}_m^H \boldsymbol{B} - \boldsymbol{g}_m^H \right\|_2 \leq \epsilon \sqrt{\frac{N_\varphi}{2\pi}}, \quad m = 1, \dots, M. \tag{3.63}$$

The number of samples N_φ has to be sufficiently large such that the Nyquist sampling criterion is satisfied for $\exp(-jM\varphi)$ and all components of $\boldsymbol{b}_{\boldsymbol{p}_{xy}}(\varphi)$.

In (3.59), the weight vectors can be optimized independently from each other. Hence, for the mth weight vector we solve

$$\begin{aligned} \min_{\boldsymbol{w}_m} \quad & \|\boldsymbol{w}_m\|_2^2 \\ \text{s.t.} \quad & \left\|\boldsymbol{g}_m - \boldsymbol{B}^H \boldsymbol{w}_m\right\|_2^2 \leq \epsilon^2 \frac{N_\varphi}{2\pi}. \end{aligned} \tag{3.64}$$

The latter problem is feasible if and only if

$$\left\|(\boldsymbol{I}_{N_\varphi} - \boldsymbol{\Pi}_{\boldsymbol{B}^H})\boldsymbol{g}_m\right\|_2 \leq \epsilon \sqrt{\frac{N_\varphi}{2\pi}}, \tag{3.65}$$

where $\boldsymbol{\Pi}_{\boldsymbol{B}^H}$ is the orthogonal projection matrix onto the column-space of $\boldsymbol{B}^H$.

The problem in (3.64) can be solved efficiently by means of a 1D search over the Lagrange parameter, similar as it has been explained in [Hud81,LSW03]. The Lagrangian of (3.64) can be written as

$$\mathcal{L}(\boldsymbol{w}_m, \lambda_m) = \|\boldsymbol{w}_m\|_2^2 + \lambda_m \left(\left\|\boldsymbol{g}_m - \boldsymbol{B}^H \boldsymbol{w}_m\right\|_2^2 - \epsilon^2 \frac{N_\varphi}{2\pi} \right), \tag{3.66}$$

where $\lambda_m \geq 0$ is the Lagrange parameter. Taking the derivative of the Lagrangian with respect to the weight vector and setting it to zero yields [Bra83]

$$\boldsymbol{w}_m(\lambda_m) = \lambda_m \left(\boldsymbol{I}_{N^2 - N} + \lambda_m \boldsymbol{B}\,\boldsymbol{B}^H\right)^{-1} \boldsymbol{B}\,\boldsymbol{g}_m. \tag{3.67}$$

The latter weight vector minimizes the Lagrangian for a fixed value of λ_m.

The rows of $\boldsymbol{B}$ form conjugate pairs. Hence, there is a unitary matrix $\boldsymbol{Q}$ such that

$$\boldsymbol{G} = \boldsymbol{Q}\boldsymbol{B} \tag{3.68}$$

is real-valued, where each row of $\boldsymbol{Q}$ contains exactly two non-zero entries. The singular value decomposition (SVD) of $\boldsymbol{G}$ can be expressed as

$$\boldsymbol{G} = \boldsymbol{U}_G \boldsymbol{\Gamma}_G \boldsymbol{V}_G^T, \tag{3.69}$$

where the diagonal matrix

$$\boldsymbol{\Gamma}_G = \operatorname{diag}\{\gamma_1, \dots, \gamma_{r_G}\}, \quad \gamma_1 \geq \gamma_2 \geq \dots \geq \gamma_{r_G} > 0 \tag{3.70}$$

contains the singular values in non-ascending order, r_G is the rank of $\boldsymbol{G}$, and

$$\begin{aligned} \boldsymbol{U}_G &\in \mathbb{R}^{N^2-N \times r_G} \\ \boldsymbol{V}_G &\in \mathbb{R}^{N_\varphi \times r_G} \end{aligned} \tag{3.71}$$

contain the orthonormal left- and right-singular vectors, respectively.

Using (3.68) and (3.69), we obtain

$$\boldsymbol{w}_m(\lambda_m) = \lambda_m \boldsymbol{Q}^H \boldsymbol{U}_G \left(\boldsymbol{I}_{r_G} + \lambda_m \boldsymbol{\Gamma}_G^2\right)^{-1} \boldsymbol{\Gamma}_G \boldsymbol{V}_G^T \boldsymbol{g}_m. \tag{3.72}$$

Hence, the norm of $\boldsymbol{w}_m(\lambda_m)$ can be evaluated efficiently as

$$\|\boldsymbol{w}_m(\lambda_m)\|_2 = \lambda_m \left(\sum_{k=1}^{r_G} \frac{|z_{m,k}|^2 \gamma_k^2}{(1+\lambda_m \gamma_k^2)^2} \right)^{1/2}, \tag{3.73}$$

where $z_{m,k}$ is the kth component of the $r_G \times 1$ vector

$$\boldsymbol{z}_m = \boldsymbol{V}_G^T \boldsymbol{g}_m. \tag{3.74}$$

To obtain the optimum Lagrange parameter $\lambda_m^\star$, we exploit that the norm of any weight vector $\boldsymbol{w}_m$, which satisfies the constraint in (3.64) with strict inequality, can be reduced without violating the constraint. Therefore, the constraint in (3.64) has to be satisfied with equality for the optimum weight vector. Consequently, $\lambda_m^\star$ is a null of

$$h_m(\lambda_m) = \left\|\boldsymbol{g}_m - \boldsymbol{B}^H \boldsymbol{w}_m(\lambda_m)\right\|_2^2 - \epsilon^2 \frac{N_\varphi}{2\pi}. \tag{3.75}$$

Equations (3.68), (3.69), and (3.72) allow simplifying (3.75) as

$$h_m(\lambda_m) = \sum_{k=1}^{r_G} \frac{|z_{m,k}|^2}{(1+\lambda_m \gamma_k^2)^2} - \zeta_m, \tag{3.76}$$

where

$$\zeta_m = \epsilon^2 \frac{N_\varphi}{2\pi} - \|(\boldsymbol{I}_{N_\varphi} - \boldsymbol{\Pi}_{\boldsymbol{B}^H})\boldsymbol{g}_m\|_2^2 \tag{3.77}$$

is non-negative, assuming that (3.64) is feasible.

From (3.67) and (3.75), we obtain

$$h_m(0) = \|\boldsymbol{g}_m\|_2^2 - \epsilon^2 \frac{N_\varphi}{2\pi} = \frac{N_\varphi}{2\pi}\left(\frac{1}{2\pi} - \epsilon^2\right) > 0. \tag{3.78}$$

Equation (3.76) shows that $h_m(\lambda_m)$ monotonically decreases for $\lambda_m \geq 0$. Furthermore,

$$\lim_{\lambda_m \to \infty} h_m(\lambda_m) = -\zeta_m \leq 0. \tag{3.79}$$

Hence, $h_m(\lambda_m)$ has a unique positive null. Since

$$h_m(\lambda_m^\star) = \sum_{k=1}^{r_G} \frac{|z_{m,k}|^2}{(1+\lambda_m^\star \gamma_k^2)^2} - \zeta_m = 0, \tag{3.80}$$

we have

$$\begin{aligned} \sum_{k=1}^{r_G} \frac{|z_{m,k}|^2}{(1+\lambda_m^\star \gamma_1^2)^2} - \zeta_m \leq 0 \quad &\Rightarrow \lambda_m^\star \geq \frac{1}{\gamma_1^2}\left(\frac{\|\boldsymbol{z}_m\|_2}{\sqrt{\zeta_m}} - 1\right) \\ \sum_{k=1}^{r_G} \frac{|z_{m,k}|^2}{(1+\lambda_m^\star \gamma_{r_G}^2)^2} - \zeta_m \geq 0 \quad &\Rightarrow \lambda_m^\star \leq \frac{1}{\gamma_{r_G}^2}\left(\frac{\|\boldsymbol{z}_m\|_2}{\sqrt{\zeta_m}} - 1\right) \\ \sum_{k=1}^{r_G} \frac{|z_{m,k}|^2}{(\lambda_m^\star \gamma_k^2)^2} - \zeta_m \geq 0 \quad &\Rightarrow \lambda_m^\star \leq \left(\frac{1}{\zeta_m}\sum_{k=1}^{r_G}\frac{|z_{m,k}|^2}{\gamma_k^4}\right)^{1/2}. \end{aligned} \tag{3.81}$$

The bounds in (3.81) limit the uncertainty in the optimum Lagrange parameter. This parameter can be computed efficiently using, for example, Newton's method. Our simulation results suggest that the complexity for computing $f_M(\boldsymbol{p}_{xy})$ is dominated by the SVD of $\boldsymbol{G}$. If $N_\varphi \geq N^2 - N$, the computational complexity of the SVD of $\boldsymbol{G}$ is $\mathcal{O}(N_\varphi N^4)$ [GvL96, TBI97]. In comparison, the computation of the number of contiguous lags for a candidate MRA geometry requires only $\mathcal{O}(N^2)$ operations. Moreover, the evaluation of $d_{\mathrm{Co}}(\boldsymbol{p}_{xy})$ for a candidate Cornwell array requires $\mathcal{O}(N^4)$ operations. Hence, the computational complexity to evaluate a candidate array geometry is significantly higher for the proposed method as compared to the MRA and Cornwell methods. Compared to the MRA methods, the proposed method has the advantage that the sensor coordinates are continuous optimization variables, and therefore continuous optimization techniques (such as sequential quadratic programming) can be used to obtain locally optimum arrays.

3.3 Fourier-domain covariance augmentation technique

Similar to the augmented covariance matrix $\boldsymbol{C}_{\mathrm{CA}}$, we define the $(M_{\mathrm{FD}}+1)\times(M_{\mathrm{FD}}+1)$ covariance matrix

$$\boldsymbol{C}_{\mathrm{FDCA}} \triangleq \sum_{l=1}^{L} P_l \boldsymbol{a}_{\mathrm{FDCA}}(\varphi_l)\boldsymbol{a}_{\mathrm{FDCA}}^H(\varphi_l) + \sigma \boldsymbol{I}_{M_{\mathrm{FD}}+1}, \tag{3.82}$$

where

$$\boldsymbol{a}_{\mathrm{FDCA}}(\varphi) \triangleq \left[1, e^{-j\varphi}, \ldots, e^{-jM_{\mathrm{FD}}\varphi}\right]^T \tag{3.83}$$

is an $(M_{\mathrm{FD}}+1)\times 1$ Vandermonde vector. Since $\boldsymbol{C}_{\mathrm{FDCA}}$ is Toeplitz and Hermitian, it is fully determined by the entries of its first column

$$[\boldsymbol{C}_{\mathrm{FDCA}}]_{m,1} = \sum_{l=1}^{L} P_l e^{-j(m-1)\varphi_l} + \sigma\delta_{m,1}, \quad m = 1, \ldots, M_{\mathrm{FD}}+1. \tag{3.84}$$

Hence, the off-diagonal entries of $\boldsymbol{C}_{\mathrm{FDCA}}$ are Fourier series coefficients of $\Psi_{\varphi}(\varphi)$ scaled by 2π.

The first entry of the first column of $\boldsymbol{C}_{\mathrm{FDCA}}$ can be estimated as

$$\left[\hat{\boldsymbol{C}}_{\mathrm{FDCA}}\right]_{1,1} = \mathrm{tr}\{\hat{\boldsymbol{R}}_{\boldsymbol{x}}\}/N, \tag{3.85}$$

where the estimation error is given by

$$\left[\hat{\boldsymbol{C}}_{\mathrm{FDCA}} - \boldsymbol{C}_{\mathrm{FDCA}}\right]_{1,1} = \mathrm{tr}\{\boldsymbol{\Delta}_{\boldsymbol{R}_{\boldsymbol{x}}}\}/N, \tag{3.86}$$

and $\mathrm{tr}\{\cdot\}$ denotes the trace operator. Using (3.44), the remaining entries of the first column of $\boldsymbol{C}_{\mathrm{FDCA}}$ can be estimated as

$$\left[\hat{\boldsymbol{C}}_{\mathrm{FDCA}}\right]_{m+1,1} = 2\pi \boldsymbol{w}_m^H \underline{\mathrm{vec}}\{\hat{\boldsymbol{R}}_{\boldsymbol{x}}\}, \quad m = 1, \ldots, M_{\mathrm{FD}}, \tag{3.87}$$

where the estimation errors are given by

$$\left[\hat{\boldsymbol{C}}_{\mathrm{FDCA}} - \boldsymbol{C}_{\mathrm{FDCA}}\right]_{m+1,1} = 2\pi \left\langle \delta_m(\varphi), \Psi_{\varphi}(\varphi)\right\rangle + 2\pi \boldsymbol{w}_m^H \underline{\mathrm{vec}}\{\boldsymbol{\Delta}_{\boldsymbol{R}_{\boldsymbol{x}}}\}, \quad m = 1, \ldots, M_{\mathrm{FD}}. \tag{3.88}$$

The essence of the FDCA technique is to estimate the DOAs by applying the root-MUSIC algorithm to the Toeplitz Hermitian matrix $\hat{\boldsymbol{C}}_{\mathrm{FDCA}}$.

The vector function $\boldsymbol{a}_{\mathrm{FDCA}}(\varphi)$ is one-to-one on $[-\pi, \pi)$, and $M_{\mathrm{FD}}+1$ pairwise different Vandermonde vectors of length $M_{\mathrm{FD}}+1$ are always linearly independent. Therefore, the azimuth angles of $L < M_{\mathrm{FD}} + 1$ sources can be computed from $\boldsymbol{C}_{\mathrm{FDCA}}$ without ambiguities for the full 360° field-of-view. Consequently, the FDCA technique allows to unambiguously estimate up to M_{FD} source DOAs, assuming that $\hat{\boldsymbol{C}}_{\mathrm{FDCA}}$ is a sufficiently accurate estimate of $\boldsymbol{C}_{\mathrm{FDCA}}$. In particular, the FDCA technique allows estimating the DOAs of more sources than there are sensors if $M_{\mathrm{FD}} > N$.

The first term on the right hand side of (3.88) leads to biased DOA estimates, as it does not depend on the received data. The DOA estimates of the FDCA technique are therefore inconsistent. Furthermore, we cannot conclude that the proposed array geometries allow identifying M_{FD} source DOAs due to the approximation errors $\delta_m(\varphi)$. However, simulation results suggest that these drawbacks are typically irrelevant in practice, because the finite sample errors usually dominate (3.88) (see Section 3.4).

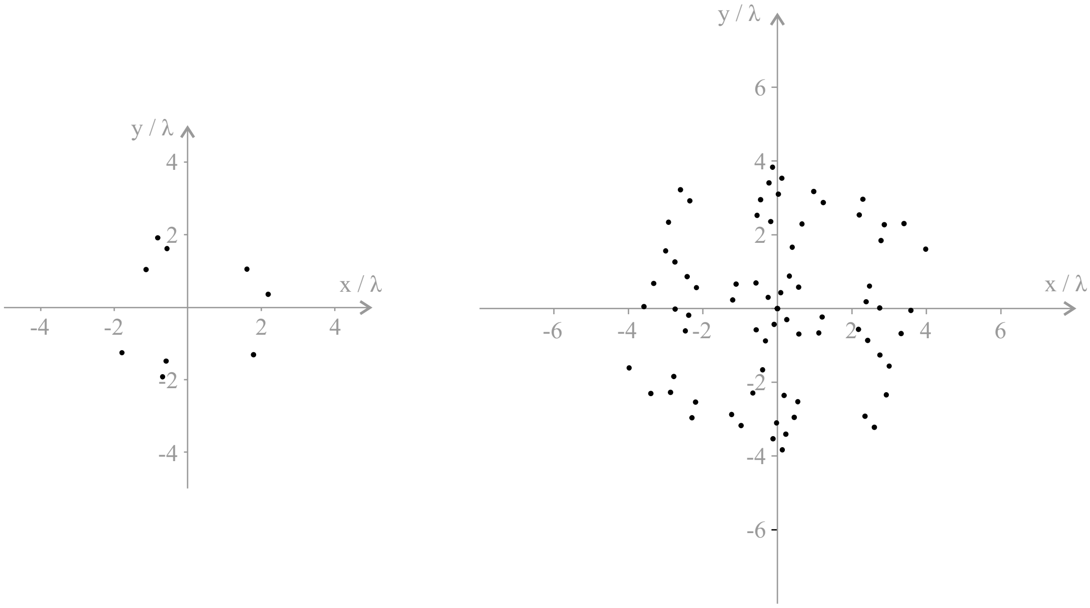
The DOA estimation MSE performance of the CA technique is often significantly worse than the stochastic CRB [AGGS98]. Our simulation results show that also the FDCA technique does not achieve the stochastic CRB. However, the DOA estimates of both the CA and FDCA techniques can be refined efficiently by a local maximization of the likelihood function as suggested in [AGGS98] in application to the CA approach.

3.4 Simulation results

To obtain good feasible points of (3.54), we generated 10^5 independent random arrays for each value of M by distributing the sensors uniformly inside a sphere of radius one. For each of these arrays, we determined the scaling of $\boldsymbol{p}_{xy}$ that minimizes $f_M(\boldsymbol{p}_{xy})$. Subsequently, we performed a local optimization of the best 10^3 scaled arrays. We chose $\epsilon = 10^{-2}/\sqrt{2\pi}$, so that the L^2-norms of the approximation errors $\delta_m(\varphi)$ are 40 dB smaller than the L^2-norms of $\exp(-jm\varphi)/(2\pi)$. Moreover, we chose $\xi = 1.25/(2\pi)$. Figures 3.7–3.10 depict the arrays obtained for $N = 9, \ldots, 12$. In Table 3.1, we list the sensor locations.

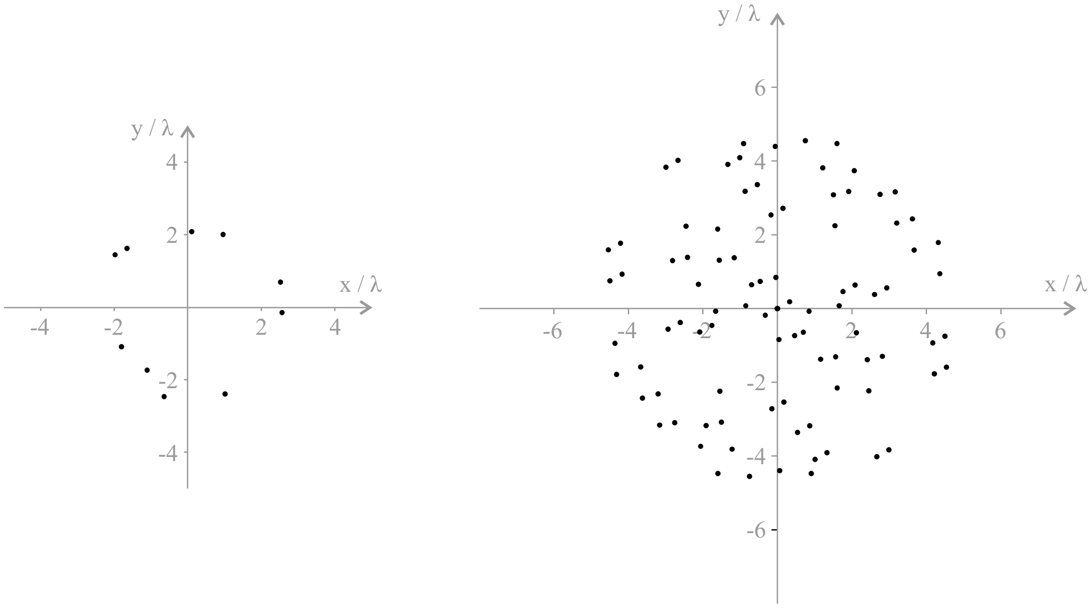
The number of randomly generated arrays was not sufficient so that the best locally optimum points of $f_M(\boldsymbol{p}_y)$ occurred several times. Hence, the arrays depicted in Figures 3.7–3.10 are probably not globally optimum points of (3.58).

The apertures of the proposed arrays are substantially larger than the apertures of standard planar MRAs with the same number of sensors. For example, the apertures of the planar MRAs depicted in Figures 3.1 and 3.2 lie between 1.5λ and $1.5\sqrt{2}\lambda$, depending on the line of sight. In contrast, Figure 3.7(a) shows that the aperture of



(a) (b)

Figure 3.7. (a) Proposed array geometry for $N = 9$ sensors. (b) Corresponding co-array.



(a) (b)

Figure 3.8. (a) Proposed array geometry for $N = 10$ sensors. (b) Corresponding co-array.

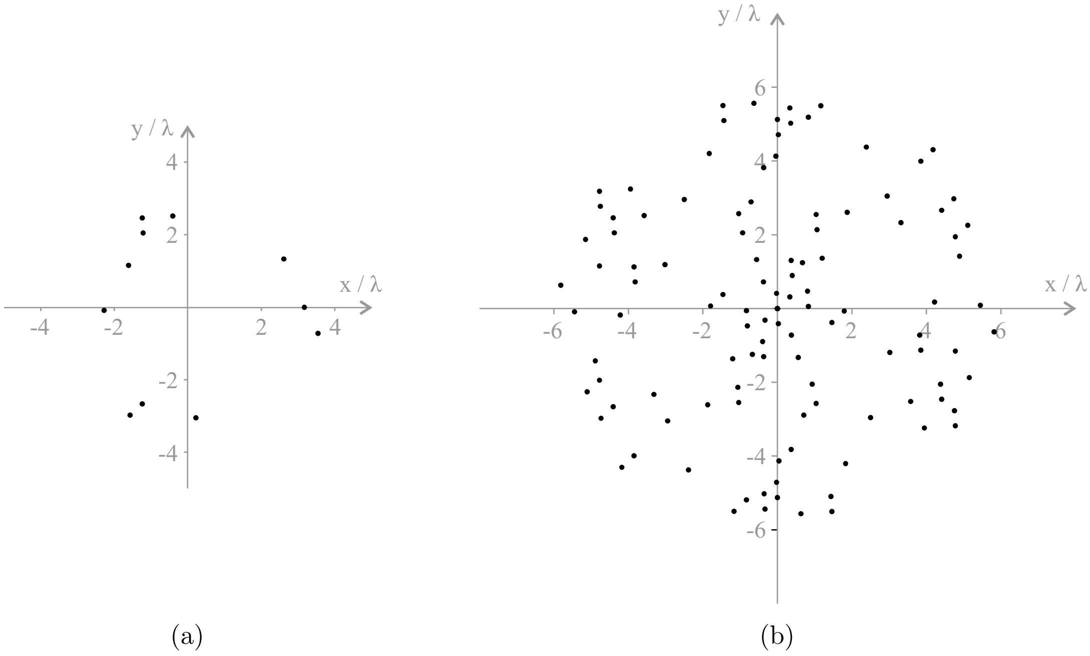


Figure 3.9. (a) Proposed array geometry for $N = 11$ sensors. (b) Corresponding co-array.

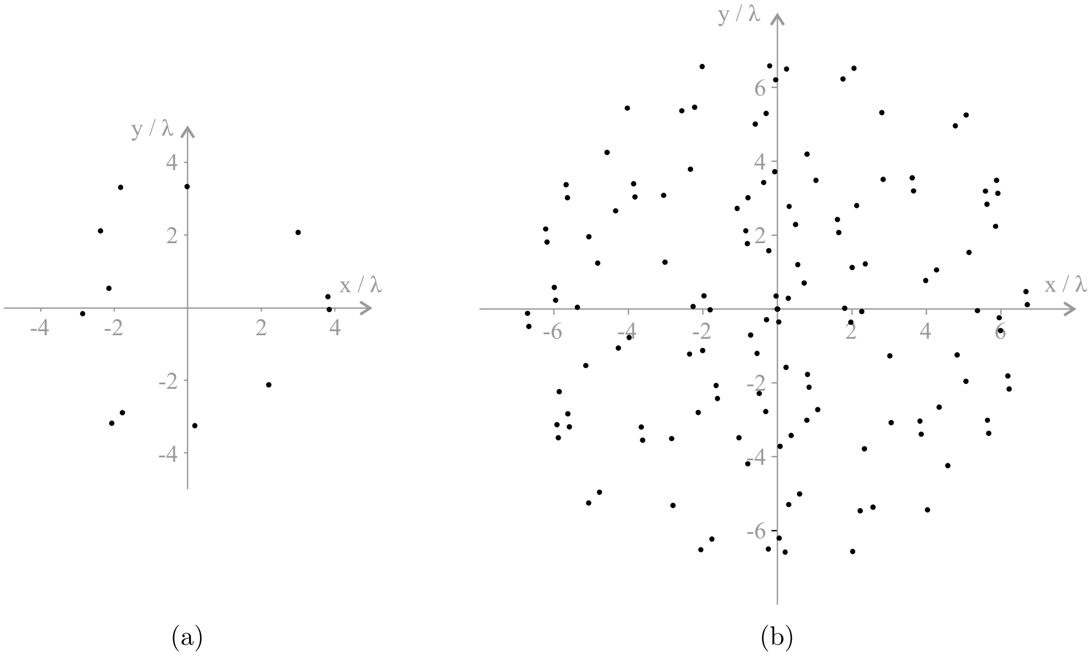


Figure 3.10. (a) Proposed array geometry for $N = 12$ sensors. (b) Corresponding co-array.

	Figure 3.7		Figure 3.8		Figure 3.9		Figure 3.10	
n	p_{x_n}/λ	p_{y_n}/λ	p_{x_n}/λ	p_{y_n}/λ	p_{x_n}/λ	p_{y_n}/λ	p_{x_n}/λ	p_{y_n}/λ
1	1.608	1.049	2.521	0.693	-0.405	2.518	-0.016	3.336
2	-1.793	-1.256	-1.650	1.623	2.611	1.328	-1.776	-2.892
3	-0.683	-1.916	1.018	-2.393	-1.568	-2.979	-1.825	3.313
4	-1.136	1.035	0.959	2.003	-1.608	1.153	3.811	0.300
5	-0.566	1.615	-1.974	1.442	-2.277	-0.089	-2.859	-0.170
6	1.789	-1.308	2.563	-0.150	0.224	-3.049	3.849	-0.051
7	-0.816	1.915	-1.800	-1.094	-1.238	-2.666	3.001	2.069
8	2.187	0.355	0.111	2.081	3.541	-0.722	-2.372	2.114
9	-0.591	-1.489	-1.106	-1.736	-1.212	2.049	2.204	-2.123
10			-0.641	-2.468	3.169	-0.001	0.195	-3.250
11					-1.238	2.458	-2.142	0.535
12							-2.069	-3.183

Table 3.1. Sensor locations of the proposed arrays.

the proposed array of $N = 9$ sensors is approximately 4λ. Due to their large apertures, the proposed arrays provide better azimuth resolution.

For the arrays of Figures 3.7–3.10, we have

$$\begin{aligned} N = 9, M_{\mathrm{FD}} = 23 \quad &\Rightarrow \tfrac{N^2}{M_{\mathrm{FD}}} \approx 3.522 \\ N = 10, M_{\mathrm{FD}} = 29 \quad &\Rightarrow \tfrac{N^2}{M_{\mathrm{FD}}} \approx 3.448 \\ N = 11, M_{\mathrm{FD}} = 35 \quad &\Rightarrow \tfrac{N^2}{M_{\mathrm{FD}}} \approx 3.457 \\ N = 12, M_{\mathrm{FD}} = 42 \quad &\Rightarrow \tfrac{N^2}{M_{\mathrm{FD}}} \approx 3.429. \end{aligned} \tag{3.89}$$

The latter results suggest that M_{FD} increases quadratically with the number of sensors. It can therefore be expected that also the apertures of the proposed arrays increase quadratically with the number of sensors.

Next, we evaluate the performance of the FDCA technique. Thereto, we consider the array of $N = 12$ sensors depicted in Figure 3.10. The following simulation results are averaged over the sources.

In our first example, we consider the case when there are more sources than sensors. Specifically, we assume that $L = 15$ equal-power signals with $\mathrm{SNR} = 0$ dB impinge on the array from the DOAs $\varphi_l = (l-1)10^\circ(\pi/180^\circ)$. In this case, it is not possible to estimate the signal DOAs by means of the conventional MUSIC algorithm. We show results for the RMSE and standard deviation of the FDCA estimates. Furthermore, we show the RMSE performance of the refined FDCA estimates, obtained by a local maximization of the likelihood function. As a reference, we also plot the stochastic CRB.

Figure 3.11 shows that the FDCA technique reliably estimates all signal DOAs, but it does not achieve the stochastic CRB. However, the RMSE performance of the refined FDCA estimates obtained by a local maximization of the likelihood function is close to the stochastic CRB. Moreover, the RMSEs and standard deviations of the FDCA estimates are close to each other. Thus, the approximation errors $\delta_m(\varphi)$ do not lead to significantly biased DOA estimates. Also, note that the FDCA technique requires rather large numbers of snapshots to resolve all sources.

In the next example, we vary the angular separation between the sources. The DOA of the lth source is $\varphi_l = (l-1)\Delta\varphi$. The number of snapshots has been set to $K = 1000$. All other parameters are chosen as before.

Figure 3.12 shows that the FDCA technique resolves all the sources if their angular separation is above 8°. The azimuth resolution capability can be improved by increasing

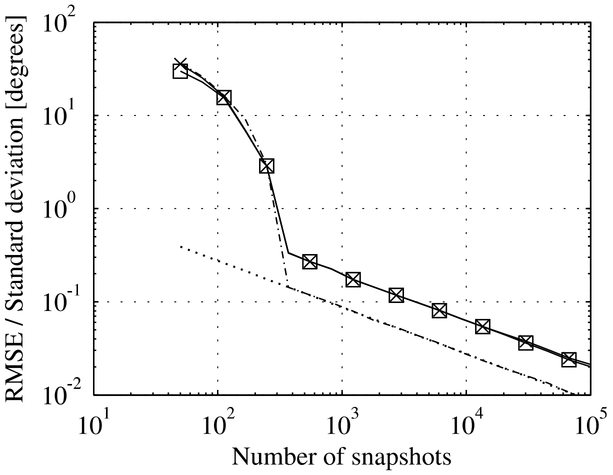


Figure 3.11. DOA estimation performance for more sources than sensors. Solid line with crosses: RMSE of the FDCA estimates. Solid line with squares: Standard deviation of the FDCA estimates. Dash-dotted line: RMSE of the refined FDCA estimates, based on a local maximization of the likelihood function. Dotted line: Stochastic CRB.

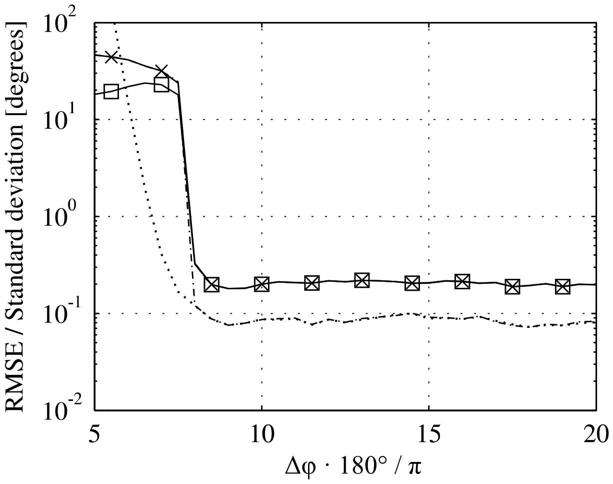


Figure 3.12. DOA estimation performance for more sources than sensors. Solid line with crosses: RMSE of the FDCA estimates. Solid line with squares: Standard deviation of the FDCA estimates. Dash-dotted line: RMSE of the refined FDCA estimates, based on a local maximization of the likelihood function. Dotted line: Stochastic CRB.

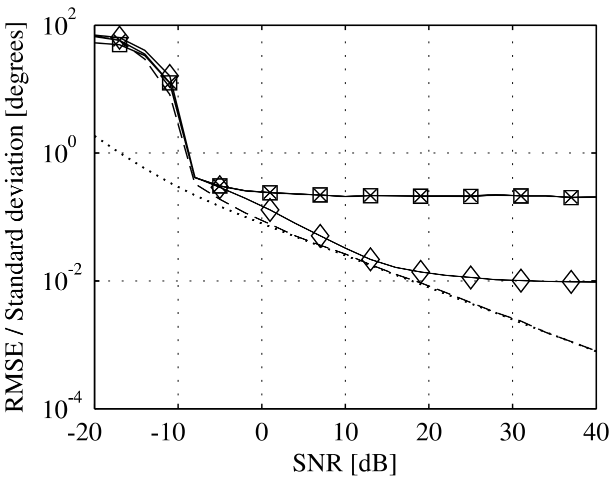


Figure 3.13. DOA estimation performance in the case of manifold ambiguities. Solid line with crosses: RMSE of the FDCA estimates. Solid line with squares: Standard deviation of the FDCA estimates. Dashed line: RMSE of the modified FD root-MUSIC technique, using compressed sensing to determine the source roots. Solid line with diamonds: RMSE of the Capon estimates. Dotted line: Stochastic CRB.

the number of sensors. Since the results in (3.89) suggest that M_{FD} increases quadratically with the number of sensors, it can be expected that a two times larger number of sensors leads to a four times better angular resolution capability.

In our third example, we consider a scenario with manifold ambiguity. We assume $L = 7$ equal-power sources with the DOAs $-176.45°$, $-132.50°$, $-24.74°$, $29.53°$, $77.57°$, $125.74°$, and $153.89°$. A manifold ambiguity occurs, because the steering vector for the DOA $173.24°$ lies approximately within $\mathcal{R}(\boldsymbol{A}(\boldsymbol{\varphi}))$. Hence, the noise-free MUSIC null-spectrum function is approximately zero for $\varphi = 173.24°\pi/180°$. Thus, choosing the L smallest minima of the MUSIC null-spectrum function leads to outliers among the DOA estimates. There are several alternatives to the FDCA technique, which allow to resolve manifold ambiguities. For example, the performance breakdown of the MUSIC algorithm due to manifold ambiguities can be avoided by determining those minima of the MUSIC null-spectrum function that correspond to the actual sources by means of methods based on the l^1-norm ("compressed sensing") [ASG99b]. Also, the Capon estimator does not suffer from manifold ambiguities [vT02]. In Figure 3.13, we compare the FDCA technique with the Capon estimator and the modified FD root-MUSIC technique, which uses the compressed sensing method of [SBL11] to select the source roots. The array covariance matrix is estimated using $K = 200$ independent snapshot vectors.

Figure 3.13 demonstrates that the FDCA technique reliably resolves the manifold ambiguity. However, Figure 3.13 also shows that the modified FD root-MUSIC technique and the Capon estimator outperform the FDCA technique. The main benefit of the FDCA technique is therefore that it allows estimating the DOAs of more sources than sensors.

3.5 Summary

We have proposed a novel approach to the design of sparse arrays for azimuthal DOA estimation. The proposed array design is related to the MRA concept, as it maximizes the number of available contiguous Fourier series coefficients of the angular power density function. The sensors are not required to lie on a uniform grid. Our simulation results suggest that the apertures of the proposed arrays increase quadratically with the number of sensors.

Based on the proposed array geometry design, we developed a subspace-based DOA estimation technique, which allows estimating the DOAs of more sources than sensors, using only second-order statistics of the received data. The proposed DOA estimation technique is related to the CA technique, but in contrast to the latter method, it provides non-ambiguous DOA estimates for the 360° azimuth field-of-view.

Chapter 4

Narrowband beamforming using multi-dimensional covariance fitting

4.1 Background

A common objective of narrowband adaptive beamforming techniques is to maximize the output signal-to-interference-plus-noise ratio (OSINR). The OSINR can be maximized by minimizing the total output power subject to the constraint that the desired signal is not distorted. This leads to the MVDR beamformer [Cap69], which provides an excellent performance and a fast convergence rate if the beamformer training data does not contain the desired signal component [RMB74]. The latter condition on the training data may be satisfied in some active radar and sonar systems, but there are numerous applications (such as wireless communications, passive location in radar and sonar, etc.), in which the training data is "contaminated" by the desired signal component. If the beamformer training data contains the desired signal component, then even small estimation errors in the signal steering vector and/or array covariance matrix may lead to a severe degradation of the MVDR beamformer performance. These estimation errors can result in a cancelation of the desired signal component, because the distortionless response constraint is formulated based on the presumed signal steering vector, without taking into account model errors. The suppression of the desired signal component due to model errors is commonly referred to as *signal self-nulling* [God97, vT02].

Various techniques have been proposed to improve the robustness of the MVDR beamformer against signal self-nulling. One such technique is to stabilize the array response towards the desired signal by specifying its value for different steering vectors in the vicinity of the presumed signal steering vector [God97, vT02]. Alternatively, the derivative of the beamformer power response towards the desired signal (and possibly also higher order derivatives) can be forced to zero. Another popular technique is to limit the beamformer sensitivity to signal steering vector estimation errors by means of a norm-constraint on the weight vector, which leads to diagonal loading-type solutions [Abr81, Hud81]. Furthermore, the beamformer robustness can be improved by applying eigenspace processing to the erroneous estimate of the signal steering vector [God97, vT02].

More recently, several beamformers based on worst-case output power minimization have been proposed [VGL03, SGLW03, VGLM04, LB05, GSS$^+$10]. The general concept of these beamformers is to minimize the maximum output power for the set of presumed array covariance matrices subject to constraints on the beamformer response for the set of presumed signal steering vectors. To simplify these optimization problems, the worst-case principle is used.

Different sets have been used to model the presumed errors in the array covariance matrix and signal steering vector. For example, in [VGL03] a spherical set has been used for the signal steering vectors, and array covariance matrix estimation errors have not been taken into account. This approach has been generalized in [LB05] and [LSW03] to ellipsoidal signal steering vector sets. Both signal steering vector and array covariance matrix estimation errors have been considered in [SGLW03] and [VGLM04]. The latter two papers use different models for the array covariance matrix estimation errors. Further, in [SGLW03] the generalization to higher-rank signal components has been studied.

In [LSW03], it has been shown that the beamformers based on worst-case output power minimization of [VGL03] and [LB05] can be formulated equivalently as a 1D covariance fitting problem. The optimum steering vector of this 1D covariance fitting problem can be interpreted as a refined estimate of the true signal steering vector, which is then used to compute the beamformer weight vector using the MVDR beamforming technique.

The method proposed in this chapter is based on the observation that the refined estimate of the signal steering vector obtained from 1D covariance fitting tends towards the principal eigenvector of the sample covariance matrix. However, this principal eigenvector represents a weighted sum of the steering vectors of all sources. Therefore, the 1D covariance fitting problem of [LSW03] leads to an inherently non-optimum refined estimate of the signal steering vector in the presence of interferers. To reduce the detrimental effect of interferers on the refined estimate of the signal steering vector, we extend the 1D covariance fitting approach to the framework of MD covariance fitting, modeling the source steering vectors by means of uncertainty sets [RGa]. The MD covariance fitting approach provides refined estimates of the source steering vectors. Subsequently, the beamformer weight vector is computed by means of the MVDR beamforming technique, using the refined estimate of the signal steering vector.

Note that the array covariance matrix allows to identify the subspace that is spanned by the steering vectors of all sources, but it does not allow to identify the source steering vectors. However, the source steering vectors can be estimated in several important applications such as cellular or satellite communication systems using pilot symbols or (approximate) knowledge of the array manifold. Estimates of the interferer steering vectors are often used in adaptive receive beamforming to improve the interference suppression by means of null steering [vT02]. In contrast, the proposed MD covariance fitting technique uses information about the interferer steering vectors to compute a refined estimate of the signal steering vector. We show that the information about the interferer steering vectors helps to reduce the uncertainty in the signal steering vector.

The proposed MD covariance fitting approach results in a non-convex optimization problem. We develop a local convex approximation of this problem, which belongs to the class of SDP problems [BV04]. Such problems can be solved using, for example, the logarithmic barrier method [NN94,BV04]. The complexities to compute the barrier function, its gradient, and its Hessian are derived.

Our simulation results demonstrate that the proposed beamformer based on MD covariance fitting achieves a substantially improved performance as compared to the current state-of-the-art robust adaptive beamformers.

4.1.1 System model

Consider L narrowband sources that impinge on an array of N sensors. The beamformer output at the kth time instant is computed as

$$y(k) = \boldsymbol{w}^H \boldsymbol{x}(k), \tag{4.1}$$

where $\boldsymbol{x}(k) \in \mathbb{C}^{N \times 1}$ and $\boldsymbol{w} \in \mathbb{C}^{N \times 1}$ are the array snapshot and beamformer weight vectors, respectively.

The narrowband snapshot vector is modeled as

$$\boldsymbol{x}(k) = \sum_{l=1}^{L} \boldsymbol{a}_l s_l(k) + \boldsymbol{n}(k) = \boldsymbol{A}\boldsymbol{s}(k) + \boldsymbol{n}(k), \tag{4.2}$$

where $\boldsymbol{a}_l \in \mathbb{C}^{N \times 1}$ is the steering vector (spatial signature) of the lth source, $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_L] \in \mathbb{C}^{N \times L}$ is the matrix of the source spatial signatures, $s_l(k)$ is the complex envelope of the lth source at the kth time instant, $\boldsymbol{s}(k) = [s_1(k), \ldots, s_L(k)]^T \in \mathbb{C}^{L \times 1}$ is the source waveform vector, and $\boldsymbol{n}(k) \in \mathbb{C}^{N \times 1}$ is the noise vector. Other than in Chapters 2 and 3, we do not require that the sources are located in the far-field, and therefore we do not specify the parametrization of the source steering vectors. In the sequel, we assume that the steering vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_L$ are linearly independent and that the signal and noise waveforms are uncorrelated zero-mean random processes. Without loss of generality, we assume that the source with index $l = 1$ is the source-of-interest, while the sources with indices $l = 2, \dots, L$ are interferers. The number of sources L is assumed to be less than the number of sensors N.

The beamformer performance is measured in terms of the OSINR, defined as [vT02]

$$\mathrm{OSINR} \triangleq \frac{P_1 |\boldsymbol{w}^H \boldsymbol{a}_1|^2}{\boldsymbol{w}^H \boldsymbol{R}_{i+n} \boldsymbol{w}}, \tag{4.3}$$

where

$$\boldsymbol{R}_{i+n} = \sum_{l=2}^{L} P_l \boldsymbol{a}_l \boldsymbol{a}_l^H + \sigma \boldsymbol{I}_N \tag{4.4}$$

is the interference-plus-noise covariance matrix, $P_l = \mathrm{E}\{|s_l(k)|^2\}$ is the power of the lth source, $\mathrm{E}\{\cdot\}$ denotes the statistical expectation, σ is the sensor noise power, and $\boldsymbol{I}_N$ stands for the $N \times N$ identity matrix.

4.1.2 Narrowband MVDR beamformer

The OSINR can be maximized by minimizing the output interference-plus-noise power subject to the constraint that the desired signal remains undistorted. This can be formulated as

$$\min_{\boldsymbol{w}} \ \boldsymbol{w}^H \boldsymbol{R}_{i+n} \boldsymbol{w} \quad \text{s.t.} \quad \boldsymbol{w}^H \boldsymbol{a}_1 = 1. \tag{4.5}$$

The solution of (4.5) is given by [vT02]

$$\boldsymbol{w}_{\mathrm{opt}} = \frac{\boldsymbol{R}_{i+n}^{-1} \boldsymbol{a}_1}{\boldsymbol{a}_1^H \boldsymbol{R}_{i+n}^{-1} \boldsymbol{a}_1}. \tag{4.6}$$

Substituting (4.6) in (4.3) yields

$$\mathrm{OSINR}_{\mathrm{opt}} = P_1 \boldsymbol{a}_1^H \boldsymbol{R}_{i+n}^{-1} \boldsymbol{a}_1. \tag{4.7}$$

The optimum weight vector of (4.5) does not change if the interference-plus-noise covariance matrix is replaced by the array covariance matrix

$$\boldsymbol{R}_{\boldsymbol{x}} = P_1 \boldsymbol{a}_1 \boldsymbol{a}_1^H + \boldsymbol{R}_{i+n} = \boldsymbol{A} \boldsymbol{P} \boldsymbol{A}^H + \sigma \boldsymbol{I}_N, \tag{4.8}$$

where the diagonal matrix $\boldsymbol{P} = \mathrm{diag}\{P_1, \dots, P_L\}$ contains the source powers. The array covariance matrix can be estimated from the received snapshot vectors as

$$\hat{\boldsymbol{R}}_{\boldsymbol{x}} = \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{x}(k) \boldsymbol{x}^H(k). \tag{4.9}$$

In the following, we assume that $\hat{\boldsymbol{R}}_{\boldsymbol{x}}$ is positive definite. Replacing $\boldsymbol{R}_{i+n}$ in (4.6) by the sample covariance matrix $\hat{\boldsymbol{R}}_{\boldsymbol{x}}$ and $\boldsymbol{a}_1$ by the estimated signal steering vector $\hat{\boldsymbol{a}}_1$ leads to the sample matrix inversion (SMI) based MVDR beamformer [vT02]

$$\boldsymbol{w}_{\text{MVDR}} = \frac{\hat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1} \hat{\boldsymbol{a}}_1}{\hat{\boldsymbol{a}}_1^H \hat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1} \hat{\boldsymbol{a}}_1}. \tag{4.10}$$

The performance of this beamformer is known to suffer severely from errors in the presumed signal steering vector and array covariance matrix if the snapshot vectors that are used in (4.9) contain the desired signal component. In this case, the performance of the MVDR beamformer can degrade severely due to signal self-nulling effects [Ger99].

Over the last three decades, multiple techniques have been proposed to improve the robustness of the MVDR beamformer against errors in the presumed signal steering vector and array covariance matrix. In the following, we briefly review several popular robust beamforming techniques.

4.1.3 Point and derivative mainlobe constraints

The array response towards the desired signal can be stabilized by specifying its value for different steering vectors in the vicinity of the presumed signal steering vector. Alternatively, the derivative (and possibly also higher-order derivatives) of the beamformer power response can be forced to zero. The derivative is taken with respect to uncertain parameters such as the DOA or the frequency of the desired signal. Both the point and derivative constraints can be formulated as linear constraints on the weight vector $\boldsymbol{w}$. The resulting linearly constrained minimum variance (LCMV) beamforming problem can be expressed as

$$\min_{\boldsymbol{w}} \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}} \boldsymbol{w} \quad \text{s.t.} \quad \boldsymbol{C}^H \boldsymbol{w} = \boldsymbol{h}, \tag{4.11}$$

where $\boldsymbol{C}$ is the $N \times J$ constraint matrix, $\boldsymbol{h}$ is the $J \times 1$ constraint vector, and $J < N$ is the number of linear constraints. If $\boldsymbol{C}$ has full-column rank, then the solution of (4.11) is given by

$$\boldsymbol{w}_{\text{LCMV}} = \hat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1} \boldsymbol{C} \left(\boldsymbol{C}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1} \boldsymbol{C} \right)^{-1} \boldsymbol{h}. \tag{4.12}$$

For example, assuming far-field sources and a linear array, $\boldsymbol{C}$ and $\boldsymbol{h}$ may be chosen as [vT02]

$$\begin{aligned} \boldsymbol{C} &= \left[\boldsymbol{a}(\hat{\theta}_1), \boldsymbol{a}(\hat{\theta}_1 - \Delta\theta), \boldsymbol{a}(\hat{\theta}_1 + \Delta\theta) \right] \\ \boldsymbol{h} &= [1, 1, 1]^T, \end{aligned} \tag{4.13}$$

where $\hat{\theta}_1$ is the presumed DOA of the desired signal, and $\Delta\theta$ corresponds to the expected DOA estimation errors. The latter choice of $\boldsymbol{C}$ and $\boldsymbol{h}$ leads to three distortionless constraints for different nearby directions. This provides robustness against DOA estimation errors. Alternatively, $\boldsymbol{C}$ and $\boldsymbol{h}$ may be chosen as [vT02]

$$\begin{aligned} \boldsymbol{C} &= \left[\boldsymbol{a}(\hat{\theta}_1), \frac{\partial \boldsymbol{a}(\hat{\theta}_1)}{\partial \theta}, \frac{\partial^2 \boldsymbol{a}(\hat{\theta}_1)}{\partial \theta^2} \right] \\ \boldsymbol{h} &= [1, 0, 0]^T \end{aligned} \tag{4.14}$$

such that the first- and second-order derivatives of the beampattern are equal to zero for the presumed DOA of the desired signal.

Note that the point and derivative constraints do not provide robustness against arbitrary-type errors in the signal steering vector. Furthermore, each linear equality constraint reduces the beamformer degrees of freedom by one.

4.1.4 Norm constraints

The beamformer response towards the desired signal is

$$\boldsymbol{w}^H \boldsymbol{a}_1 = \boldsymbol{w}^H (\hat{\boldsymbol{a}}_1 - \boldsymbol{\delta}_1), \tag{4.15}$$

where $\boldsymbol{\delta}_1$ is the steering vector estimation error. Based on (4.15), we define the beamformer sensitivity to signal steering vector estimation errors as

$$T_{\text{se}} \triangleq \left\| \frac{\boldsymbol{w}}{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1} \right\|_2^2, \tag{4.16}$$

where $\|\cdot\|_2$ denotes the Euclidean vector norm. The smaller T_{se}, the more robust is the beamformer with respect to signal steering vector estimation errors. The sensitivity has been defined in [GM55, CZO87, vT02] without the normalization factor $1/(\boldsymbol{w}^H \hat{\boldsymbol{a}}_1)$, because most of the traditional adaptive beamformers use a distortionless constraint such that $\boldsymbol{w}^H \hat{\boldsymbol{a}}_1 = 1$. However, we will see shortly that some modern beamformers do not use any distortionless constraint. Therefore, to compare the sensitivity of such beamformers, it is important to normalize the beamformer weight vector. Such normalization is immaterial for the OSINR performance.

Applying the Cauchy-Schwarz inequality to the denominator of (4.16) shows that

$$T_{\text{se}} \geq \frac{1}{\|\hat{\boldsymbol{a}}_1\|_2^2}. \tag{4.17}$$

A straightforward approach to limit the beamformer sensitivity is to enforce a distortionless constraint and an upper bound on the norm of the beamformer weight vector. This leads to the norm-bounded MVDR beamformer [Hud81]

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}} \boldsymbol{w} \\ \text{s.t.} \quad & \boldsymbol{w}^H \hat{\boldsymbol{a}}_1 = 1 \\ & \boldsymbol{w}^H \boldsymbol{w} \leq \alpha_w, \end{aligned} \tag{4.18}$$

where α_w is a user-defined parameter. The sensitivity of this beamformer to signal steering vector estimation errors is at most α_w. It can be seen easily that

$$\alpha_w \geq \frac{1}{\|\hat{\boldsymbol{a}}_1\|_2^2} \tag{4.19}$$

is a necessary and sufficient condition for the feasibility of (4.18).

Equations (4.3) and (4.6) give

$$\text{OSINR}_{\text{opt}} \leq \frac{P_1}{\sigma \|\boldsymbol{w}_{\text{opt}}\|_2^2}. \tag{4.20}$$

Applying the Cauchy-Schwarz inequality to the distortionless constraint in (4.5) yields

$$\|\boldsymbol{w}_{\text{opt}}\|_2 \|\boldsymbol{a}_1\|_2 \geq 1. \tag{4.21}$$

The inequalities in (4.20) and (4.21) yield

$$\text{OSINR}_{\text{opt}} \leq \frac{P_1}{\sigma} \|\boldsymbol{a}_1\|_2^2. \tag{4.22}$$

In many scenarios of practical interest, in particular if there are few interferers that are spatially well-separated from the source of the desired signal, we have

$$\text{OSINR}_{\text{opt}} \approx \frac{P_1}{\sigma} \|\boldsymbol{a}_1\|_2^2. \tag{4.23}$$

Then, (4.21) is approximately satisfied with equality, i.e., $\|\boldsymbol{w}_{\text{opt}}\|_2$ is only slightly larger than its lower bound $1/\|\boldsymbol{a}_1\|_2$. In the presence of model errors, the norm of the beamformer weight vector should be smaller than in the error-free case. Therefore, choosing α_w moderately (1–3 dB) above its lower bound in (4.19) often leads to good performance results.

The problem in (4.18) belongs to the class of second-order-cone programming (SOCP) problems [BV04]. This problem can be solved as follows. If the MVDR beamformer weight vector (4.10) satisfies the norm-constraint in (4.18), then this weight vector solves (4.18). Otherwise, the optimum weight vector of (4.18) has to satisfy the norm-constraint with equality. In this case, the optimum weight vector can be expressed as [Hud81]

$$\boldsymbol{w}(\eta) = \frac{(\hat{\boldsymbol{R}}_{\boldsymbol{x}} + \eta \boldsymbol{I}_N)^{-1}\hat{\boldsymbol{a}}_1}{\hat{\boldsymbol{a}}_1^H(\hat{\boldsymbol{R}}_{\boldsymbol{x}} + \eta \boldsymbol{I}_N)^{-1}\hat{\boldsymbol{a}}_1}, \tag{4.24}$$

where η is a positive diagonal loading factor. It has been shown in [LSW04] that $f(\eta) = \|\boldsymbol{w}(\eta)\|_2^2$ is a monotonically decreasing function. The point at which $f(\eta) = \alpha_w$ can be determined efficiently using, for example, Newton's method. Further details on the efficient implementation are presented in [LSW03, LSW04]. The computational complexity for solving (4.18) is $\mathcal{O}(N^3)$.

4.1.5 Eigenspace projections

The essence of the eigenspace projection technique is to improve the erroneous estimate of the signal steering vector by projecting it onto the subspace that is spanned by the source steering vectors [CY92].

The projection of the presumed signal steering vector onto the estimated signal-and-interference subspace can be computed as

$$\boldsymbol{a}_{\mathrm{EV}} = \hat{\boldsymbol{V}}_{\boldsymbol{x},S}\hat{\boldsymbol{V}}_{\boldsymbol{x},S}^H\hat{\boldsymbol{a}}_1, \tag{4.25}$$

where $\hat{\boldsymbol{V}}_{\boldsymbol{x},S}$ has been defined in (2.13). Then, the beamformer weight vector is obtained by replacing $\hat{\boldsymbol{a}}_1$ by $\boldsymbol{a}_{\mathrm{EV}}$ in (4.10). The eigenvector projection technique is known to substantially improve the robustness for moderate values of the SNR, in particular if $N \gg L$. However, knowledge of the number of sources is required, and for low SNR values signal subspace swaps can lead to a severe performance degradation [TST95, HNS01].

4.1.6 Worst-case output power minimization

More recently, several beamformers based on worst-case output power minimization have been proposed [VGL03, SGLW03, VGLM04, LB05]. For example, the beamformer of [VGL03] minimizes the total output power subject to the constraint that the beamformer power response is larger than one for all steering vectors within a sphere that is centered on the presumed signal steering vector $\hat{\boldsymbol{a}}_1$. This can be formulated as

$$\min_{\boldsymbol{w}} \ \boldsymbol{w}^H\hat{\boldsymbol{R}}_{\boldsymbol{x}}\boldsymbol{w} \quad \text{s.t.} \quad |\boldsymbol{w}^H(\hat{\boldsymbol{a}}_1 + \tilde{\boldsymbol{\delta}})| \geq 1 \quad \forall\ \|\tilde{\boldsymbol{\delta}}\|_2 \leq \epsilon_1, \tag{4.26}$$

where ϵ_1 is a user-defined parameter. Similar to the beamformer in (4.18), the beamformer in (4.26) leads to diagonal loading-type beamformer weight vectors with diagonal loading factors that depend on $\hat{\boldsymbol{R}}_{\boldsymbol{x}}$, $\hat{\boldsymbol{a}}_1$, and ϵ_1 in a non-trivial way [VGL03]. In particular, the beamformers (4.18) and (4.26) lead to the same performance if ϵ_1 and α_w are chosen appropriately.

The constraints in (4.26) are satisfied if and only if they are satisfied for the worst-case mismatch vector. That means,

$$\min_{\tilde{\boldsymbol{\delta}}} |\boldsymbol{w}^H(\hat{\boldsymbol{a}}_1 + \tilde{\boldsymbol{\delta}})| \quad \text{s.t.} \quad \|\tilde{\boldsymbol{\delta}}\|_2 \leq \epsilon_1 \tag{4.27}$$

has to be greater than or equal to one. If $\epsilon_1 < \|\hat{\boldsymbol{a}}_1\|_2$, then the solution of (4.27) is given by [VGL03]

$$|\boldsymbol{w}^H \hat{\boldsymbol{a}}_1| - \epsilon_1 \|\boldsymbol{w}\|_2, \tag{4.28}$$

which is attained for the worst-case mismatch vector

$$\boldsymbol{\delta}_{1,\text{wc}} = -\frac{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1}{\|\boldsymbol{w}^H \hat{\boldsymbol{a}}_1\|_2} \frac{\boldsymbol{w}}{\|\boldsymbol{w}\|_2} \epsilon_1. \tag{4.29}$$

Hence, (4.26) can be formulated equivalently as

$$\min_{\boldsymbol{w}} \; \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}} \boldsymbol{w} \quad \text{s.t.} \quad |\boldsymbol{w}^H \hat{\boldsymbol{a}}_1| - \epsilon_1 \|\boldsymbol{w}\|_2 \geq 1. \tag{4.30}$$

The objective and constraint functions in (4.30) are invariant with respect to the multiplication of $\boldsymbol{w}$ with a unit-modulus complex number. Hence, the solution of (4.30) is equal to the solution of

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}} \boldsymbol{w} \\ \text{s.t.} \quad & \Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1\right\} - \epsilon_1 \|\boldsymbol{w}\|_2 \geq 1. \end{aligned} \tag{4.31}$$

The problem (4.31) is a SOCP problem [BV04].

Note that the beamformer (4.26) minimizes the worst-case output power, but this does not imply that this beamformer maximizes the worst-case OSINR. The maximization of the worst-case OSINR is possible only if an estimate of the interference-plus-noise covariance matrix $\boldsymbol{R}_{\text{i+n}}$ is available. However, we assume here that such an estimate is not available.

The robustness of (4.31) to signal steering vector estimation errors can be explained as follows. The constraint in (4.30) gives

$$T_{\text{se}} \leq \frac{\|\boldsymbol{w}\|_2^2}{(1 + \epsilon_1 \|\boldsymbol{w}\|_2)^2} \leq \frac{1}{\epsilon_1^2}, \tag{4.32}$$

so this constraint strictly limits the beamformer sensitivity. Moreover, (4.15) yields

$$|\boldsymbol{w}^H \boldsymbol{a}_1| \geq \Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1\right\} - \|\boldsymbol{\delta}_1\|_2 \, \|\boldsymbol{w}\|_2. \tag{4.33}$$

The constraint in (4.31) therefore implies that

$$|\boldsymbol{w}^H \boldsymbol{a}_1| \geq 1 + (\epsilon_1 - \|\boldsymbol{\delta}_1\|_2) \, \|\boldsymbol{w}\|_2. \tag{4.34}$$

The minimization of the beamformer output power leads to a small gain $|\boldsymbol{w}^H \boldsymbol{a}_1|^2$ towards the desired signal. The larger the power of the desired signal, the stronger is the incentive for a small value of $|\boldsymbol{w}^H \boldsymbol{a}_1|^2$. Due to (4.34), this results in small values of $\|\boldsymbol{w}\|_2$ if the signal steering vector lies within the presumed uncertainty set, i.e., if

$$\|\boldsymbol{\delta}_1\|_2 < \epsilon_1. \tag{4.35}$$

The upper bound on the sensitivity in (4.32) is a strictly increasing function of $\|\boldsymbol{w}\|_2$. Consequently, the constraint of the beamformer (4.31) leads to an incentive for a low sensitivity to signal steering vector estimation errors. The strength of this incentive increases with the power of the desired signal. A similar signal-dependent incentive for a low sensitivity does not exist for the norm-bounded MVDR beamformer.

Several generalizations of (4.26) have been discussed in the literature. For example, in [SWL03, LSW03, LB05], and [LLGS10], $\tilde{\boldsymbol{\delta}}$ is confined to an ellipsoidal set, which provides more flexibility in modeling the signal steering vector estimation errors. Note that spherical and ellipsoidal steering vector uncertainty sets both lead to second-order-cone constraints [BV04], which can be implemented very similarly in convex optimization solvers. For the ease of notation, we do not further consider ellipsoidal steering vector sets, even though the generalization to these sets is straightforward.

It has been shown in [LSW03] that the problem in (4.31) can be formulated equivalently as

$$\begin{aligned} \max_{\tilde{\boldsymbol{a}}_1, \tilde{P}_1} \quad & \tilde{P}_1 \\ \text{s.t.} \quad & \hat{\boldsymbol{R}}_{\boldsymbol{x}} \succeq \tilde{P}_1 \tilde{\boldsymbol{a}}_1 \tilde{\boldsymbol{a}}_1^H \\ & \|\tilde{\boldsymbol{a}}_1 - \hat{\boldsymbol{a}}_1\|_2 \leq \epsilon_1, \end{aligned} \tag{4.36}$$

where $\succeq$ is the matrix inequality for Hermitian matrices, and tildes are used to indicate the optimization variables. Note that (4.36) can be interpreted as a 1D covariance fitting problem. The unique optimum steering vector of (4.36) can be computed with the complexity $\mathcal{O}(N^3)$ [LSW03]. The beamformer weight vector of the 1D covariance fitting problem is computed based on the MVDR concept as

$$\boldsymbol{w}_{1\mathrm{D}} = \frac{\hat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1} \boldsymbol{a}_{1\mathrm{D}}}{\boldsymbol{a}_{1\mathrm{D}}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1} \boldsymbol{a}_{1\mathrm{D}}}, \tag{4.37}$$

where $\boldsymbol{a}_{\text{1D}}$ is the optimum steering vector of (4.36). It has been shown in [LSW03] that $\boldsymbol{w}_{\text{1D}}$ solves (4.31).

It can be verified numerically that $\boldsymbol{a}_{\text{1D}}$ is in general not collinear to $\boldsymbol{a}_1$, even when the true values $\boldsymbol{R}_{\boldsymbol{x}}$ and $\boldsymbol{a}_1$ are used in (4.36) instead of $\hat{\boldsymbol{R}}_{\boldsymbol{x}}$ and $\hat{\boldsymbol{a}}_1$, respectively. This becomes intuitively clear from (4.36), as $\tilde{\boldsymbol{a}}_1$ tends towards the principal eigenvector of $\hat{\boldsymbol{R}}_{\boldsymbol{x}}$. However, this eigenvector represents a weighted sum of the steering vectors of all sources, and therefore the vector $\tilde{\boldsymbol{a}}_1$ tends towards the interferer steering vectors. This undesirable effect can be attributed to the 1D nature of the covariance fitting approach in (4.36), which does not take into account the presence of interferers.

4.2 Multi-dimensional covariance fitting

To reduce the detrimental effect of interferers on the refined estimate of the signal steering vector, we extend in this section the 1D covariance fitting approach to MD covariance fitting. Towards this end, we take into account information about the interferer steering vectors. Specifically, we assume that $\hat{\boldsymbol{a}}_l$ is an estimate of the steering vector of the lth source, and ϵ_l is an estimated upper bound on the norm of $\hat{\boldsymbol{a}}_l - \boldsymbol{a}_l$. Note that in practice, such estimates can be available at the array. For example, in wireless communication systems such estimates can be obtained using pilot sequences.

The proposed MD covariance fitting technique determines refined estimates of the steering vectors of all sources. Subsequently, the beamformer weight vector is computed based on the MVDR concept, using the refined estimate of the signal steering vector.

Information about the interferer steering vectors is often used in adaptive receive beamformers to improve the suppression of interferers. Thereto, a popular technique is to enforce the null-steering constraints [vT02]

$$\boldsymbol{w}^H \hat{\boldsymbol{a}}_l = 0, \quad l = 2, \dots, L. \tag{4.38}$$

All beamformers discussed above and also the MD covariance fitting technique proposed in this section can be combined straightforwardly with null steering. For notational simplicity, we introduce all beamformers without taking into account null steering. In our simulations, we compare the beamformer performance with and without null steering.

The proposed MD covariance fitting approach can be formulated as

$$\begin{aligned} \max_{\tilde{\boldsymbol{A}},\tilde{\boldsymbol{P}},\tilde{\sigma}\geq 0} \quad & \log\det\left(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N\right) \\ \text{s.t.} \quad & \hat{\boldsymbol{R}}_{\boldsymbol{x}} \succeq \tilde{\boldsymbol{A}}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N \\ & \|(\tilde{\boldsymbol{A}} - \hat{\boldsymbol{A}})\boldsymbol{e}_l\|_2 \leq \epsilon_l, \quad l = 1, \ldots, L \\ & \tilde{\boldsymbol{P}} = \tilde{\boldsymbol{P}} \odot \boldsymbol{I}_L \succeq \boldsymbol{0}, \end{aligned} \tag{4.39}$$

where $\hat{\boldsymbol{A}} = [\hat{\boldsymbol{a}}_1, \ldots, \hat{\boldsymbol{a}}_L]$ is the estimated signal steering matrix, $\odot$ is the Schur-Hadamard element-wise matrix product, and $\boldsymbol{e}_l$ is a column-vector of conformable dimension whose lth entry is equal to one and all other entries are equal to zero. The constraint in the last line of (4.39) ensures that the $L \times L$ matrix $\tilde{\boldsymbol{P}}$ is positive semidefinite and diagonal. The steering vector uncertainty sets are assumed to be sufficiently separated one from another so that the columns of $\tilde{\boldsymbol{A}}$ are pairwise linearly independent. The latter condition prevents the permutation ambiguity among the columns of $\tilde{\boldsymbol{A}}$.

Note that $\det(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N)$ is proportional to the volume of the ellipsoid

$$\mathcal{E}_1 = \{\boldsymbol{x} | \boldsymbol{x}^H(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N)^{-1}\boldsymbol{x} \leq 1\}. \tag{4.40}$$

The constraint in the second line of (4.39) can be written equivalently as

$$\hat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1} \preceq \left(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N\right)^{-1}. \tag{4.41}$$

Hence, each $\boldsymbol{x}$ within $\mathcal{E}_1$ also lies within the ellipsoid

$$\mathcal{E}_2 = \{\boldsymbol{x} | \boldsymbol{x}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1} \boldsymbol{x} \leq 1\}. \tag{4.42}$$

The MD covariance fitting is therefore achieved by maximizing the volume of $\mathcal{E}_1$ subject to the constraint that $\mathcal{E}_1$ fits within $\mathcal{E}_2$.

4.2.1 Optimality of the true parameters

Proposition 4.1. *If* $\hat{\boldsymbol{R}}_{\boldsymbol{x}} = \boldsymbol{R}_{\boldsymbol{x}}$ *and* $\|\hat{\boldsymbol{a}}_l - \boldsymbol{a}_l\|_2 \leq \epsilon_l$ *for* $l = 1, \ldots, L$*, then* $(\boldsymbol{A}, \boldsymbol{P}, \sigma)$ *is an optimum point of* (4.39)*.*

Proof. It can be verified easily that $(\boldsymbol{A}, \boldsymbol{P}, \sigma)$ is a feasible point. The value of the objective function for this point is $\log\det \boldsymbol{R}_{\boldsymbol{x}}$. The objective function is upper bounded by $\log\det \boldsymbol{R}_{\boldsymbol{x}}$ due to the constraint in the second line of (4.39). $\square$

Note that the similar statement does not hold for the 1D covariance fitting problem. In the presence of interferers, the signal steering vector $\boldsymbol{a}_1$ is not a solution of (4.36), even if $\hat{\boldsymbol{R}}_{\boldsymbol{x}} = \boldsymbol{R}_{\boldsymbol{x}}$ *and* $\hat{\boldsymbol{a}}_1 = \boldsymbol{a}_1$.

4.2.2 Scaling ambiguity and its elimination

Substituting $\tilde{\boldsymbol{A}} = [\tilde{\boldsymbol{a}}_1, \ldots, \tilde{\boldsymbol{a}}_L]$ and $\tilde{\boldsymbol{P}} = \operatorname{diag}\{\tilde{P}_1, \ldots, \tilde{P}_L\}$ in (4.39) yields

$$\begin{aligned} \max_{\{\tilde{\boldsymbol{a}}_l\}, \{\tilde{P}_l\}, \tilde{\sigma} \geq 0} \quad & \log\det\left(\textstyle\sum_{l=1}^L \tilde{P}_l \tilde{\boldsymbol{a}}_l \tilde{\boldsymbol{a}}_l^H + \tilde{\sigma}\boldsymbol{I}_N\right) \\ \text{s.t.} \quad & \hat{\boldsymbol{R}}_{\boldsymbol{x}} \succeq \textstyle\sum_{l=1}^L \tilde{P}_l \tilde{\boldsymbol{a}}_l \tilde{\boldsymbol{a}}_l^H + \tilde{\sigma}\boldsymbol{I}_N \\ & \|\tilde{\boldsymbol{a}}_l - \hat{\boldsymbol{a}}_l\|_2 \leq \epsilon_l, \quad l = 1, \ldots, L \\ & \tilde{P}_l \geq 0, \quad l = 1, \ldots, L. \end{aligned} \tag{4.43}$$

The problem formulation in (4.43) shows that the vectors $\tilde{\boldsymbol{a}}_l$ are subject to a scaling ambiguity, because

$$\tilde{P}_l \tilde{\boldsymbol{a}}_l \tilde{\boldsymbol{a}}_l^H = (\tilde{P}_l |\kappa_l|^2)(\tilde{\boldsymbol{a}}_l/\kappa_l)(\tilde{\boldsymbol{a}}_l/\kappa_l)^H \tag{4.44}$$

for any complex-valued κ_l, and the constraints in the third line of (4.43) do not specify κ_l uniquely. However, the scaling ambiguity has no effect on the OSINR performance, and can therefore be considered as immaterial. In the following, we say that two vectors are *essentially equal* if they are equal up to a scaling factor. We use $\boldsymbol{v}_1 \doteq \boldsymbol{v}_2$ to express the essential equality between two vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$.

This scaling ambiguity can be eliminated from (4.43) as follows. Substituting $\breve{\boldsymbol{a}}_l = \alpha_l \tilde{\boldsymbol{a}}_l \sqrt{\tilde{P}_l}$ with $|\alpha_l| = 1$ in (4.43) yields

$$\begin{aligned} \max_{\{\breve{\boldsymbol{a}}_l\}, \{\tilde{P}_l\}, \{\alpha_l\}, \tilde{\sigma} \geq 0} \quad & \log\det\left(\textstyle\sum_{l=1}^L \breve{\boldsymbol{a}}_l \breve{\boldsymbol{a}}_l^H + \tilde{\sigma}\boldsymbol{I}_N\right) \\ \text{s.t.} \quad & \hat{\boldsymbol{R}}_{\boldsymbol{x}} \succeq \textstyle\sum_{l=1}^L \breve{\boldsymbol{a}}_l \breve{\boldsymbol{a}}_l^H + \tilde{\sigma}\boldsymbol{I}_N \\ & \left\|\breve{\boldsymbol{a}}_l/(\alpha_l\sqrt{\tilde{P}_l}) - \hat{\boldsymbol{a}}_l\right\|_2 \leq \epsilon_l, \quad l = 1, \ldots, L \\ & \tilde{P}_l \geq 0, \quad l = 1, \ldots, L \\ & |\alpha_l| = 1, \quad l = 1, \ldots, L. \end{aligned} \tag{4.45}$$

The objective and the first constraint functions in (4.45) do not depend on the variables $\tilde{P}_l$ and α_l, $l = 1, \ldots, L$. These optimization variables can be chosen to satisfy the constraints in the third line of (4.45). We can therefore replace the constraints in the last three lines of (4.45) by

$$\min_{\gamma_l} \left\|\gamma_l \breve{\boldsymbol{a}}_l - \hat{\boldsymbol{a}}_l\right\|_2 \leq \epsilon_l, \quad l = 1, \ldots, L. \tag{4.46}$$

The minimum in (4.46) is attained for $\gamma_l = \breve{\boldsymbol{a}}_l^H \hat{\boldsymbol{a}}_l \,/\, \|\breve{\boldsymbol{a}}_l\|_2^2$. Using this result, (4.46) can be written as

$$|\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{a}}_l| \geq \sqrt{\|\hat{\boldsymbol{a}}_l\|_2^2 - \epsilon_l^2}\, \|\breve{\boldsymbol{a}}_l\|_2, \quad l = 1, \ldots, L. \tag{4.47}$$

The latter constraints are invariant with respect to the multiplication of $\breve{\boldsymbol{a}}_l$ by a unit-modulus complex number. The same invariance holds for the objective and the first constraint functions in (4.43). Thus, the solution of (4.43) does not change if we introduce the constraints

$$\Re\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{a}}_l\} = |\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{a}}_l|, \quad l = 1, \ldots, L. \tag{4.48}$$

The constraints in (4.48) eliminate the ambiguity of $\breve{\boldsymbol{a}}_l$, $l = 1, \ldots, L$ with respect to the multiplication with unit-modulus complex numbers. The problem in (4.45) can be written equivalently as

$$\begin{aligned} \max_{\breve{\boldsymbol{A}}, \tilde{\sigma} \geq 0} \quad & \log\det\left(\breve{\boldsymbol{A}}\breve{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N\right) \\ \text{s.t.} \quad & \hat{\boldsymbol{R}}_{\boldsymbol{x}} \succeq \breve{\boldsymbol{A}}\breve{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N \\ & \Re\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}}\boldsymbol{e}_l\} \geq \nu_l \, \|\breve{\boldsymbol{A}}\boldsymbol{e}_l\|_2, \quad l = 1, \ldots, L \\ & \Im\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}}\boldsymbol{e}_l\} = 0, \quad l = 1, \ldots, L, \end{aligned} \tag{4.49}$$

where $\breve{\boldsymbol{A}} = [\breve{\boldsymbol{a}}_1, \ldots, \breve{\boldsymbol{a}}_L]$,

$$\nu_l = \sqrt{\|\hat{\boldsymbol{a}}_l\|_2^2 - \epsilon_l^2}, \quad l = 1, \ldots, L, \tag{4.50}$$

and $\Im\{\cdot\}$ denotes the imaginary part operator.

Let $(\breve{\boldsymbol{A}}_{\text{MD}}, \sigma_{\text{MD}})$ denote a solution of (4.49). The constraints in the third and fourth lines of (4.49) prevent the scaling ambiguity of the columns of $\breve{\boldsymbol{A}}_{\text{MD}}$ with respect to unit-modulus complex numbers. To show that there is also no scaling ambiguity with respect to positive scaling factors, let us consider the problem

$$\begin{aligned} \max_{\breve{\boldsymbol{P}}, \tilde{\sigma} \geq 0} \quad & \log\det\left(\breve{\boldsymbol{A}}_{\text{MD}}\breve{\boldsymbol{P}}\breve{\boldsymbol{A}}_{\text{MD}}^H + \tilde{\sigma}\boldsymbol{I}_N\right) \\ \text{s.t.} \quad & \hat{\boldsymbol{R}}_{\boldsymbol{x}} \succeq \breve{\boldsymbol{A}}_{\text{MD}}\breve{\boldsymbol{P}}\breve{\boldsymbol{A}}_{\text{MD}}^H + \tilde{\sigma}\boldsymbol{I}_N \\ & \Re\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}}_{\text{MD}}\breve{\boldsymbol{P}}^{1/2}\boldsymbol{e}_l\} \geq \nu_l \, \|\breve{\boldsymbol{A}}_{\text{MD}}\breve{\boldsymbol{P}}^{1/2}\boldsymbol{e}_l\|_2, \quad l = 1, \ldots, L \\ & \Im\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}}_{\text{MD}}\breve{\boldsymbol{P}}^{1/2}\boldsymbol{e}_l\} = 0, \quad l = 1, \ldots, L \\ & \breve{\boldsymbol{P}} = \breve{\boldsymbol{P}} \odot \boldsymbol{I}_L \succeq \boldsymbol{0}, \end{aligned} \tag{4.51}$$

where the diagonal matrix $\breve{\boldsymbol{P}}$ scales the columns of $\breve{\boldsymbol{A}}_{\text{MD}}$. The constraints of the optimization problem in (4.51) are all convex, and the objective function is strictly concave with respect to $(\breve{\boldsymbol{P}}, \tilde{\sigma})$. Hence, (4.51) has a unique solution $(\breve{\boldsymbol{P}}, \tilde{\sigma}) = (\boldsymbol{I}_L, \sigma_{\text{MD}})$. Consequently, the columns of $\breve{\boldsymbol{A}}_{\text{MD}}$ are not subject to any scaling ambiguity.

4.2.3 Uniqueness

Proposition 4.2. *The optimum matrix* $\breve{\boldsymbol{A}}_{\mathrm{MD}}$ *of* (4.49) *may not be unique.*

Proof. See Appendix A.3. ☐

That (4.49) can have a non-unique solution may appear as a major disadvantage of the MD covariance fitting approach. However, in the next subsection we show that the non-uniqueness of (4.49) results from the non-identifiability of the source steering vectors. Hence, the non-uniqueness of (4.49) does not result from a non-optimum problem formulation, but from the fundamental source identifiability problem.

4.2.4 Identifiability

Proposition 4.3. *The knowledge of the array covariance matrix and that the* l*th source steering vector lies within the set*

$$\{\tilde{\boldsymbol{a}}_l \mid \|\tilde{\boldsymbol{a}}_l - \boldsymbol{a}_l\|_2 \leq \epsilon_l\} \tag{4.52}$$

with $\epsilon_l > 0$ *for* $l = 1, \ldots, L$ *is not sufficient to essentially identify the source steering vectors, i.e., to identify these vectors up to scaling factors.*

Proof. See Appendix A.4. ☐

The proofs of Proposition 4.2 and Proposition 4.3 show that the non-uniqueness of (4.49) follows from the essential non-identifiability of the source steering vectors. This non-identifiability is closely related to the fact that the information about the array covariance matrix does not allow to separate the sources [BAMCM97, Car98].

As the 1D covariance fitting problem exploits less information than the MD covariance fitting problem, the *a-priori* information of the 1D covariance fitting problem does not allow to essentially identify the signal steering vector. Despite that, the 1D covariance fitting problem has a unique optimum steering vector $\boldsymbol{a}_{1\mathrm{D}}$, i.e., the solution of this problem does not reflect the essential non-identifiability of the signal steering vector.

Proposition 4.4. *If the array covariance matrix and the interferer steering vectors are known exactly, then the signal steering vector is essentially identifiable. If also* $\|\hat{\boldsymbol{a}}_1 - \boldsymbol{a}_1\|_2 \leq \epsilon_1$*, then the unique optimum point of* (4.49) *is given by*

$$(\breve{\boldsymbol{A}}_{\mathrm{MD}}, \sigma_{\mathrm{MD}}) = (\boldsymbol{A}\boldsymbol{P}^{1/2}\boldsymbol{D}(\hat{\boldsymbol{A}}, \boldsymbol{A}), \sigma), \tag{4.53}$$

where

$$\boldsymbol{D}(\hat{\boldsymbol{A}}, \boldsymbol{A}) = \mathrm{diag}\{e^{-j\rho_1}, \ldots, e^{-j\rho_L}\}, \tag{4.54}$$

and ρ_l *is the complex phase of* $\boldsymbol{e}_l^T \hat{\boldsymbol{A}}^H \boldsymbol{A}\boldsymbol{e}_l$*.*

Proof. See Appendix A.5. ☐

Note that for the specific assumptions of Proposition 4.4, $\rho_2 = \ldots = \rho_L = 0$.

Proposition 4.4 demonstrates that the information about the interferer steering vectors helps to identify the signal steering vector in the essential sense. Furthermore, the beamformer based on MD covariance fitting achieves the optimum OSINR under the conditions of Proposition 4.4. As a result, we can expect that the beamformer based on MD covariance fitting achieves an improved performance as compared to the beamformer based on 1D covariance fitting.

4.2.5 Local convex approximation

The constraint in the second line of (4.49) can be written as [vT02]

$$\hat{\boldsymbol{R}}_{\boldsymbol{x}} \succeq \breve{\boldsymbol{A}}\breve{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N \Longleftrightarrow \begin{bmatrix} \hat{\boldsymbol{R}}_{\boldsymbol{x}} - \tilde{\sigma}\boldsymbol{I}_N & \breve{\boldsymbol{A}} \\ \breve{\boldsymbol{A}}^H & \boldsymbol{I}_L \end{bmatrix} \succeq \boldsymbol{0}. \tag{4.55}$$

Hence, this constraint is convex. Defining

$$\boldsymbol{F}(\breve{\boldsymbol{A}}, \tilde{\sigma}) = \begin{bmatrix} \hat{\boldsymbol{R}}_{\boldsymbol{x}} - \tilde{\sigma}\boldsymbol{I}_N & \breve{\boldsymbol{A}} \\ \breve{\boldsymbol{A}}^H & \boldsymbol{I}_L \end{bmatrix} \tag{4.56}$$

allows to write (4.49) as

$$\begin{aligned} \max_{\breve{\boldsymbol{A}}, \tilde{\sigma} \geq 0} \quad & \log\det\left(\breve{\boldsymbol{A}}\breve{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N\right) \\ \text{s.t.} \quad & \boldsymbol{F}(\breve{\boldsymbol{A}}, \tilde{\sigma}) \succeq \boldsymbol{0} \\ & \Re\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}}\boldsymbol{e}_l\} \geq \nu_l \, \|\breve{\boldsymbol{A}}\boldsymbol{e}_l\|_2, \quad l = 1, \ldots, L \\ & \Im\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}}\boldsymbol{e}_l\} = 0, \quad l = 1, \ldots, L. \end{aligned} \tag{4.57}$$

The constraints in (4.57) are all convex, but the objective function is non-concave. To obtain a convex optimization problem, we approximate the objective function

$$g(\breve{\boldsymbol{A}}, \tilde{\sigma}) = \log\det\left(\breve{\boldsymbol{A}}\breve{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N\right) \tag{4.58}$$

by its affine Taylor approximation. Note that there are more accurate concave approximations of (4.58) compared to the affine Taylor approximation. For example, the quadratic Taylor approximation may be used, dropping the convex component of the Hessian. Moreover, (4.58) can be approximated using the difference of convex programming technique [Tuy95,RG10a]. For the sake of simplicity, we consider in the following the affine Taylor approximation.

The affine Taylor approximation of (4.58) about $(\breve{\boldsymbol{A}}_0, \sigma_0)$ can be expressed as

$$\begin{aligned} g(\breve{\boldsymbol{A}}, \tilde{\sigma}) \approx g(\breve{\boldsymbol{A}}_0, \sigma_0) &+ \left.\frac{\partial g(\breve{\boldsymbol{A}}, \tilde{\sigma})}{\partial \tilde{\sigma}}\right|_{\substack{\breve{\boldsymbol{A}}=\breve{\boldsymbol{A}}_0 \\ \tilde{\sigma}=\sigma_0}} (\tilde{\sigma} - \sigma_0) \\ &+ \operatorname{tr}\left\{\left(\left.\frac{\partial g(\breve{\boldsymbol{A}}, \tilde{\sigma})}{\partial \Re\{\breve{\boldsymbol{A}}\}}\right|_{\substack{\breve{\boldsymbol{A}}=\breve{\boldsymbol{A}}_0 \\ \tilde{\sigma}=\sigma_0}}\right)^T \Re\left\{\breve{\boldsymbol{A}} - \breve{\boldsymbol{A}}_0\right\}\right\} \\ &+ \operatorname{tr}\left\{\left(\left.\frac{\partial g(\breve{\boldsymbol{A}}, \tilde{\sigma})}{\partial \Im\{\breve{\boldsymbol{A}}\}}\right|_{\substack{\breve{\boldsymbol{A}}=\breve{\boldsymbol{A}}_0 \\ \tilde{\sigma}=\sigma_0}}\right)^T \Im\left\{\breve{\boldsymbol{A}} - \breve{\boldsymbol{A}}_0\right\}\right\}, \end{aligned} \tag{4.59}$$

where $\operatorname{tr}\{\cdot\}$ denotes the trace operator, and the derivative of a scalar with respect to a matrix is defined element-wise, i.e.,

$$\frac{\partial g(\breve{\boldsymbol{A}}, \tilde{\sigma})}{\partial \Re\{\breve{\boldsymbol{A}}\}} \tag{4.60}$$

is an $N \times L$ matrix, which contains in its (n, l)th entry the derivative with respect to the (n, l)th component of $\Re\{\breve{\boldsymbol{A}}\}$.

We use

$$\begin{aligned} \breve{\boldsymbol{A}}_0 &= \hat{\boldsymbol{A}}\boldsymbol{P}_0 \\ \sigma_0 &= \frac{1}{N-L}\sum_{n=L+1}^{N} \hat{\lambda}_n, \end{aligned} \tag{4.61}$$

where $\boldsymbol{P}_0 = \operatorname{diag}\{\hat{P}_1, \ldots, \hat{P}_L\}$, $\hat{P}_l$ is an estimate of the power of the lth source, and $\hat{\lambda}_1, \ldots, \hat{\lambda}_N$ are the eigenvalues of $\hat{\boldsymbol{R}}_{\boldsymbol{x}}$, sorted in non-ascending order. Any of the methods discussed in Section 4.1 can be used to estimate the source powers. In our simulations, we use the norm-bounded MVDR beamformer (4.18) to estimate these powers.

Using the definition for the complex derivative suggested in [Bra83], we have

$$\begin{aligned}\frac{\partial g(\breve{\boldsymbol{A}},\tilde{\sigma})}{\partial \Re\{\breve{\boldsymbol{A}}\}} &= 2\Re\left\{\frac{\partial g(\breve{\boldsymbol{A}},\tilde{\sigma})}{\partial \breve{\boldsymbol{A}}^*}\right\}\\ \frac{\partial g(\breve{\boldsymbol{A}},\tilde{\sigma})}{\partial \Im\{\breve{\boldsymbol{A}}\}} &= 2\Im\left\{\frac{\partial g(\breve{\boldsymbol{A}},\tilde{\sigma})}{\partial \breve{\boldsymbol{A}}^*}\right\}.\end{aligned} \tag{4.62}$$

Substituting (4.62) in (4.59), we have

$$\begin{aligned} g(\breve{\boldsymbol{A}},\tilde{\sigma}) \approx\ & g(\breve{\boldsymbol{A}}_0,\sigma_0) + \left.\frac{\partial g(\breve{\boldsymbol{A}},\tilde{\sigma})}{\partial \tilde{\sigma}}\right|_{\substack{\breve{\boldsymbol{A}}=\breve{\boldsymbol{A}}_0\\ \tilde{\sigma}=\sigma_0}} (\tilde{\sigma}-\sigma_0)\\ &+ 2\Re\left\{\operatorname{tr}\left\{\left(\left.\frac{\partial g(\breve{\boldsymbol{A}},\tilde{\sigma})}{\partial \breve{\boldsymbol{A}}^*}\right|_{\substack{\breve{\boldsymbol{A}}=\breve{\boldsymbol{A}}_0\\ \tilde{\sigma}=\sigma_0}}\right)^H (\breve{\boldsymbol{A}}-\breve{\boldsymbol{A}}_0)\right\}\right\},\end{aligned} \tag{4.63}$$

To compute the derivatives in (4.63), we use that [vT02]

$$\frac{\partial \log\det \boldsymbol{M}(\mu)}{\partial \mu} = \operatorname{tr}\left\{\boldsymbol{M}^{-1}(\mu)\frac{\partial \boldsymbol{M}(\mu)}{\partial \mu}\right\}, \tag{4.64}$$

where $\boldsymbol{M}(\mu)$ is a matrix-valued positive definite differentiable function. The latter result gives

$$\begin{aligned}\frac{\partial g(\breve{\boldsymbol{A}},\tilde{\sigma})}{\partial \tilde{\sigma}} &= \operatorname{tr}\left\{\left(\breve{\boldsymbol{A}}\breve{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N\right)^{-1}\right\}\\ \frac{\partial g(\breve{\boldsymbol{A}},\tilde{\sigma})}{\partial \breve{\boldsymbol{A}}^*} &= \left(\breve{\boldsymbol{A}}\breve{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N\right)^{-1}\breve{\boldsymbol{A}}.\end{aligned} \tag{4.65}$$

The affine Taylor approximation of the objective function of the MD covariance fitting problem can therefore be formulated as

$$\begin{aligned} g(\breve{\boldsymbol{A}},\tilde{\sigma}) \approx\ & g(\breve{\boldsymbol{A}}_0,\sigma_0) + \operatorname{tr}\left\{\left(\breve{\boldsymbol{A}}_0\breve{\boldsymbol{A}}_0^H + \sigma_0\boldsymbol{I}_N\right)^{-1}\right\}(\tilde{\sigma}-\sigma_0)\\ &+ 2\Re\left\{\operatorname{tr}\left\{\breve{\boldsymbol{A}}_0^H\left(\breve{\boldsymbol{A}}_0\breve{\boldsymbol{A}}_0^H + \sigma_0\boldsymbol{I}_N\right)^{-1}(\breve{\boldsymbol{A}}-\breve{\boldsymbol{A}}_0)\right\}\right\}.\end{aligned} \tag{4.66}$$

Using (4.66), we approximate the MD covariance fitting problem (4.57) as

$$\begin{aligned} \max_{\breve{\boldsymbol{A}},\tilde{\sigma}\geq 0}\quad & 2\Re\left\{\operatorname{tr}\left\{\breve{\boldsymbol{A}}_0^H\boldsymbol{R}_{\boldsymbol{x},0}^{-1}\breve{\boldsymbol{A}}\right\}\right\} + \operatorname{tr}\left\{\boldsymbol{R}_{\boldsymbol{x},0}^{-1}\right\}\tilde{\sigma}\\ \text{s.t.}\quad & \boldsymbol{F}(\breve{\boldsymbol{A}},\tilde{\sigma}) \succeq \boldsymbol{0}\\ & \Re\{\hat{\boldsymbol{a}}_l^H\breve{\boldsymbol{A}}\boldsymbol{e}_l\} \geq \nu_l\,\|\breve{\boldsymbol{A}}\boldsymbol{e}_l\|_2,\quad l=1,\ldots,L\\ & \Im\{\hat{\boldsymbol{a}}_l^H\breve{\boldsymbol{A}}\boldsymbol{e}_l\} = 0,\quad l=1,\ldots,L,\end{aligned} \tag{4.67}$$

where $\boldsymbol{R}_{\boldsymbol{x},0} = \breve{\boldsymbol{A}}_0 \breve{\boldsymbol{A}}_0^H + \sigma_0 \boldsymbol{I}_N$.

To converge to a locally optimum point of (4.57), (4.67) may be applied iteratively. To ensure that the objective function increases in each iteration, a line search can be employed similar as in standard descent methods [BV04]. The optimum point of this line search is used as the expansion point of the Taylor series of (4.58) in the next iteration. In our simulations, the iterative application of (4.67) led to minor performance improvements. We therefore propose to solve (4.67) only once. Subsequently, the beamformer weight vector is computed by replacing $\hat{\boldsymbol{a}}_1$ by $\breve{\boldsymbol{A}}_1 \boldsymbol{e}_1$ in (4.10), where $(\breve{\boldsymbol{A}}_1, \sigma_1)$ is the optimum point of (4.67).

The problem (4.67) is an SDP problem, which can be solved, for example, by means of the logarithmic barrier method. In Appendix A.6, we show that the complexity for computing the barrier function and its first- and second-order derivatives is $\mathcal{O}(N^3 + L^2N^2)$. The complexity for computing an exact Newton step is $\mathcal{O}(L^3N^3)$ due to the inversion of the Karush-Kuhn-Tucker (KKT) matrix. The number of Newton iterations (including backtracking) required to solve (4.67) weakly depends on the problem size, and typically lies between 10 and 100, see [BV04]. Hence, the complexity of (4.67) is higher than the complexities of the beamformers discussed in Section 4.1, in particular if the number of sources is large. To solve (4.67), we use `CVX` [GB08a,GB08b], a package for solving convex optimization problems.

4.3 Simulation results

In this section, we numerically evaluate the performance of the proposed beamformer based on MD covariance fitting. We assume three far-field sources that impinge on an array of $N = 10$ sensors. The source steering vectors correspond to a ULA of omnidirectional identical sensors, which are located on the z-axis of the coordinate system with half a wavelength interelement spacing. Unless specified otherwise, the source elevation angles are $\theta_1 = 8°$, $\theta_2 = 0°$, and $\theta_3 = -30°$ (see Figure 2.1 for the definition of the coordinate system). The estimated steering vectors of the three sources are generated randomly as

$$\hat{\boldsymbol{a}}_l = \boldsymbol{a}_l + \boldsymbol{\delta}_l, \tag{4.68}$$

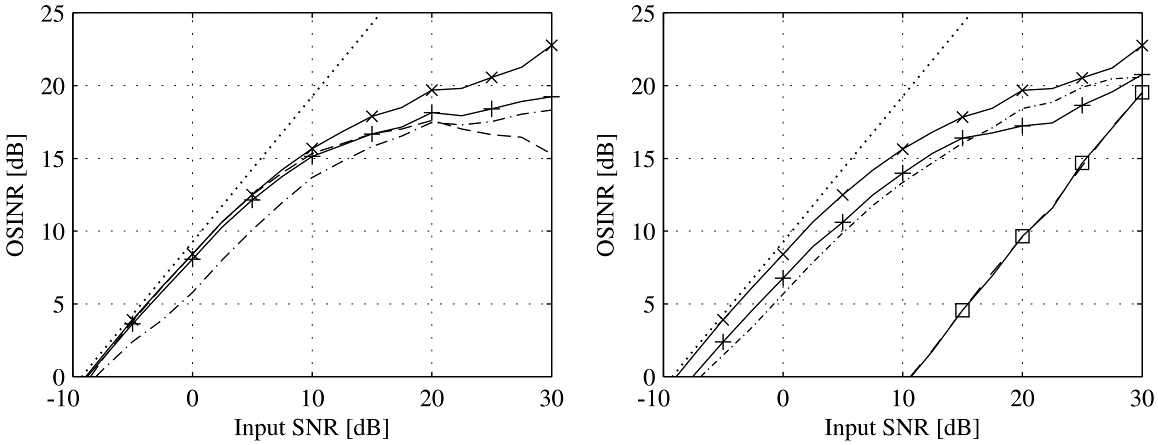
where $\boldsymbol{\delta}_l \in \mathbb{C}^{N\times 1}$ is drawn in each Monte-Carlo run independently from a zero-mean circularly symmetric complex Gaussian distribution with covariance $\mu_l \boldsymbol{I}_N$. The parameter μ_l is chosen such that $\|\boldsymbol{\delta}_l\|_2 \leq 0.5$ with probability 0.95. The bounds on the norms of the steering vector estimation errors are chosen as $\epsilon_1 = 1$ and $\epsilon_2 = \epsilon_3 = 0.5$. Unless specified otherwise, the input SNR is 20 dB, and the array covariance matrix is estimated using $K = 100$ independent snapshot vectors. The input interference-to-noise ratios (INRs) of the second and third sources are set to 40 dB and 20 dB, respectively.

We compare the performance of the proposed beamformer with the performance of the conventional non-adaptive beamformer ($\boldsymbol{w} = \hat{\boldsymbol{a}}_1$), the eigenvector projection beamformer, the norm-bounded MVDR beamformer, and the beamformer based on 1D covariance fitting. Note that the latter beamformer is equivalent to the beamformer based on worst-case output power minimization of [VGL03]. The norm-bounded MVDR beamformer uses $\alpha_w = 2/N$. As a reference, we also plot the optimum OSINR.

We also show the performance of all beamformers with extra null-steering step. The modification of the conventional beamformer and the norm-bounded MVDR beamformer is straightforward. For the eigenvector projection beamformer, we first project all source steering vectors onto the subspace spanned by the columns of $\hat{\boldsymbol{V}}_{\boldsymbol{x},S}$. Subsequently, we solve the LCMV beamforming problem taking into account the distortionless and the null-steering constraints, using the projected source steering vectors. For the beamformer based on 1D covariance fitting, we resort to the equivalent formulation (4.31), to which we add the null-steering constraints (4.38). For the beamformer based on MD covariance fitting, we first solve (4.67). Subsequently, we solve the LCMV beamforming problem with null-steering constraints, using the refined estimates of the source steering vectors.

In our first example, we study the beamformer OSINR versus the input SNR. In Figure 4.1(a), we compare the performance of the beamformers without null-steering. In this case, the performance of the beamformers based on 1D and MD covariance fitting is close to the optimum OSINR if the input SNR is below approximately 0 dB. A significant gap between the performance of these beamformers and the optimum OSINR occurs for SNRs above 0 dB. In this regime, the proposed beamformer based on MD covariance fitting outperforms the other beamformers tested. These improvements are especially pronounced for high SNRs, where signal self-nulling is prevailing. The performance of the conventional beamformer is below the OSINR range plotted.

Figure 4.1(b) depicts the performance of the beamformers with null-steering. Again, the proposed beamformer based on MD covariance fitting outperforms the other beamformers tested. In Figure 4.1(b), the curves for the conventional beamformer and the eigenvector projection beamformer overlap. The comparison with Figure 4.1(a) shows that null-steering leads to a severe performance degradation for the eigenvector projection beamformer. In contrast, null-steering leads to tremendous performance improvements for the conventional beamformer.

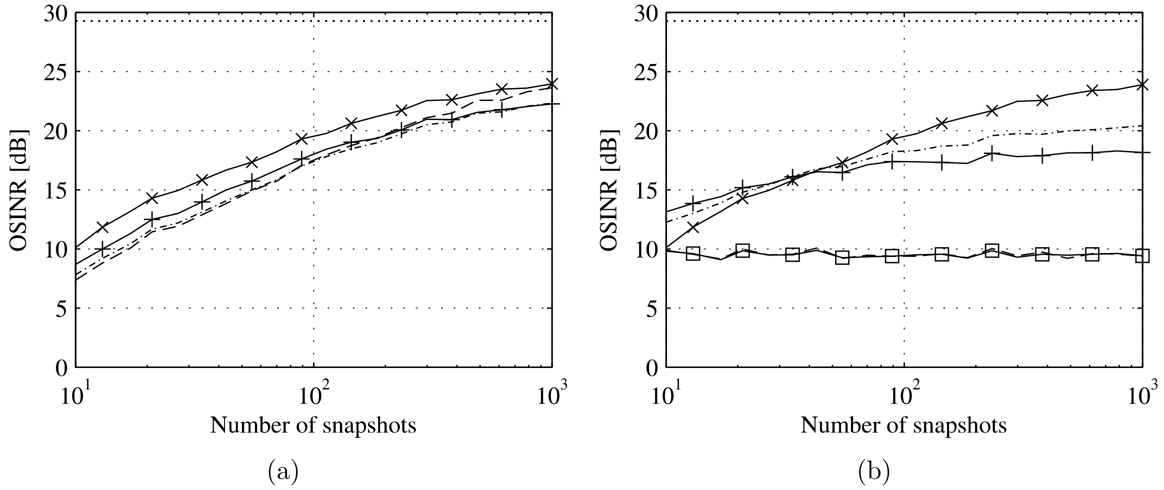


(a) (b)

Figure 4.1. Beamformer performance versus the input SNR. (a) Without null-steering. (b) With null-steering. Solid line with squares: Conventional beamformer. Dashed line: Eigenvector projection beamformer. Dash-dotted line: Norm-bounded MVDR beamformer. Solid line with pluses: Beamformer based on 1D covariance fitting. Solid line with crosses: Beamformer based on MD covariance fitting. Dotted line: Optimum OSINR.

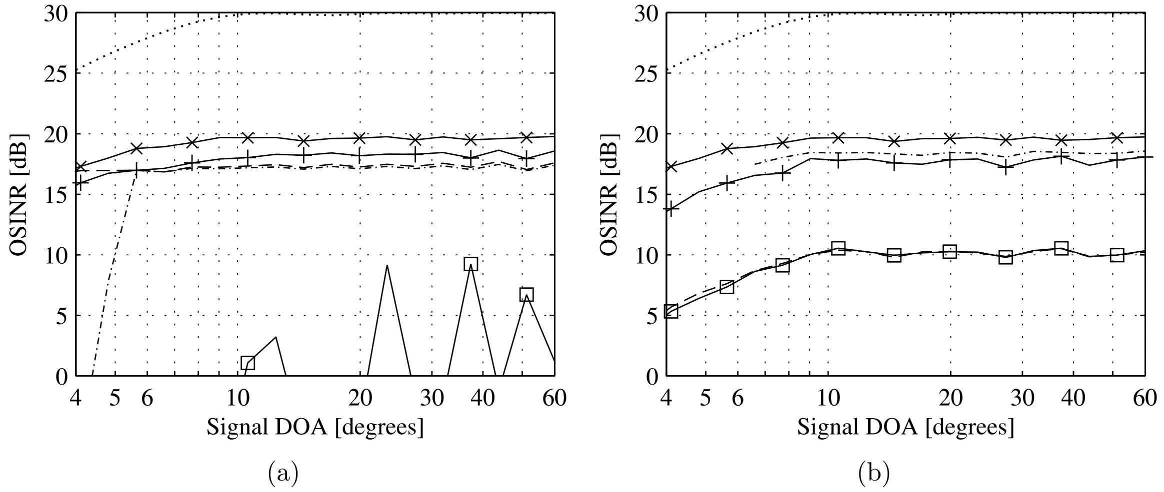
In our second example, we study the beamformer performance versus the number of snapshots that are used to estimate the array covariance matrix. The input SNR is set to 20 dB. All other parameters are chosen as before. Figure 4.2(a) shows that, if null-steering is not applied, then the beamformer based on MD covariance fitting outperforms all other beamformers. In contrast, if null-steering is used (see Figure 4.2(b)), then the beamformer based on MD covariance fitting outperforms the other beamformers only if the number of snapshots is large. Note that for adaptive beamformers, the signal steering vector should be adapted to the sample covariance matrix since otherwise signal self-nulling occurs. However, the interferer steering vectors should only be adapted to the sample covariance matrix if this matrix is a sufficiently accurate estimate of the array covariance matrix. If the number of snapshots is small, then the sample covariance matrix contains significant estimation errors. In this case, it may be unfavorable to adapt the interferer steering vectors to the sample covariance matrix. That the norm-bounded MVDR and the 1D covariance fitting based beamformers outperform the beamformer based on MD covariance fitting for small numbers of snapshots can be explained by the fact that the former two beamformers do not adapt the interferer steering vectors to the sample covariance matrix.

In our third example, we vary the DOA of the desired signal. The number of snapshots is $K = 100$. All other parameters are chosen as before. Figure 4.3 shows that the proposed beamformer again outperforms the other beamformers tested. The curve



(a) (b)

Figure 4.2. Beamformer performance versus the number of snapshots. (a) Without null-steering. (b) With null-steering. Solid line with squares: Conventional beamformer. Dashed line: Eigenvector projection beamformer. Dash-dotted line: Norm-bounded MVDR beamformer. Solid line with pluses: Beamformer based on 1D covariance fitting. Solid line with crosses: Beamformer based on MD covariance fitting. Dotted line: Optimum OSINR.



(a) (b)

Figure 4.3. Beamformer performance versus the signal DOA. (a) Without null-steering. (b) With null-steering. Solid line with squares: Conventional beamformer. Dashed line: Eigenvector projection beamformer. Dash-dotted line: Norm-bounded MVDR beamformer. Solid line with pluses: Beamformer based on 1D covariance fitting. Solid line with crosses: Beamformer based on MD covariance fitting. Dotted line: Optimum OSINR.

for the norm-bounded MVDR beamformer with null-steering does not extend to $\theta_1 = 4°\pi/180°$, because for small angular separations between the desired signal and the first interferer, the norm-bounded MVDR beamforming problem with null-steering constraints was infeasible in some Monte-Carlo runs.

4.4 Summary

We proposed a beamformer based on MD covariance fitting, which takes into account information about the interferer steering vectors to obtain an improved estimate of the signal steering vector. Simulation results show that the proposed MD covariance fitting approach leads to an improved performance as compared to the current state-of-the-art robust adaptive beamformers. The proposed beamformer involves solving an SDP problem, which can be implemented, for example, by means of the logarithmic barrier method. The complexity of computing the barrier function and its first- and second-order derivatives is $\mathcal{O}(N^3 + N^2L^2)$.

Chapter 5

Broadband beamforming based on worst-case output power minimization

5.1 Background

In this chapter, we study beamforming techniques for broadband signals with the objective to recover the waveform of the source-of-interest.

One popular approach to broadband beamforming is to decompose the baseband waveforms into narrowband frequency components by means of a DFT. Subsequently, the subbands are processed independently from each other using narrowband beamforming techniques. Hence, the methods discussed in Chapter 4 can be used to solve the narrowband beamforming problems. The time-domain beamformer output samples are obtained by applying an inverse discrete Fourier transform (IDFT) to the output samples of the narrowband beamformers. For example, DFT beamformers using narrowband techniques based on worst-case output power minimization have been studied in [WLS$^+$04]. The results of DFT beamformers are non-optimum, since correlations between the frequency domain snapshot vectors of different subbands are not taken into account. Such correlations can be reduced by increasing the DFT length [Com88].

Another approach to broadband beamforming without subband decomposition is the FIR beamformer developed by Frost [Fro72]. This beamformer uses a presteering delay front-end to time-align the desired signal components in the different sensors. The presteering delays are followed by FIR filters, and the beamformer output is obtained by summing the filtered waveforms. The filter coefficients of the broadband MVDR beamformer minimize the output power subject to a distortionless constraint for the desired signal. The latter constraint is formulated assuming that the desired signal components in the different sensors are identical after the presteering stage. However, in practice certain errors are inevitable due to, for example, signal wavefront distortions, look direction errors, array imperfections, quantization errors in presteering delays, etc. Similar as in the narrowband case, the performance of the broadband MVDR beamformer suffers severely from such model errors, because if the desired signal components are non-identical after the presteering stage, these components may be suppressed by the subsequent adaptive filter-and-sum stage.

Several methods have been proposed to improve the robustness of the broadband MVDR beamformer against signal self-nulling. These methods are often similar to the methods used to improve the robustness of the narrowband MVDR beamformer. For example, the beamformer power response towards the desired signal can be stabilized by forcing its derivative to zero. The derivative is taken, for example, with respect to the DOA of the desired signal, or with respect to the presteering delays [EC83, Tse92, TCL93, TCL95, ZT02]. However, such derivative constraints do not provide robustness against arbitrary-type signal steering vector estimation errors, and it is hardly possible to control the amount of robustness. Another popular approach to improve the robustness of the broadband MVDR beamformer is to limit the norm of the beamformer weight vector, which leads to diagonal loading-type beamformer weight vectors.

More recently, FIR beamformers based on worst-case output power minimization have been proposed [EKKG05b, EKKG05a, EKKG06]. These beamformers enforce constraints on the frequency response for sets of presumed (erroneous) signal steering vectors. As in the narrowband case, the worst-case principle is used to simplify these constraints.

In this chapter, we propose two FIR beamformers based on worst-case output power minimization [RG08b, REKGK10]. The proposed FIR beamformers use different constraints to control the frequency response towards the desired signal. The constraints of the first proposed FIR beamformer ensure that the real part of the product of the beamformer frequency response and the complex conjugate of the desired distortionless frequency response is above some lower bound for all presumed signal steering vectors. The constraints of the second proposed FIR beamformer directly limit the deviation of the beamformer frequency response from the desired distortionless frequency response for all presumed signal steering vectors.

We show that the constraints of the proposed FIR beamformers strictly limit the sensitivity to signal steering vector estimation errors, similar to the constraints of the norm-bounded broadband MVDR beamformer. Furthermore, the constraints of the proposed FIR beamformers lead to an incentive for a low sensitivity if the signal steering vector estimation errors are not underestimated. This incentive becomes stronger with increasing signal powers. Such an incentive does not exist for the norm-bounded broadband MVDR beamformer.

Even though the proposed FIR beamformers use different constraints to control the frequency response towards the desired signal, we show that these beamformers are closely related. In fact, adding a distortionless constraint to both beamformers results in the same beamformer. We also analyze the relation between the beamformer that results from adding a distortionless constraint to the proposed FIR beamformers and the norm-bounded broadband MVDR beamformer. Hence, we establish a relation between beamforming based on worst-case output power minimization and diagonal loading for the FIR beamformer framework.

The proposed FIR beamformers enforce constraints only for a finite set of frequencies. This frequency discretization leads to SOCP beamforming problems with computational complexities that are comparable to the complexity of the norm-bounded broadband MVDR beamformer. Based on the theory of positive trigonometric polynomials, we also develop a modified version of the first proposed FIR beamformer, which avoids the frequency discretization and associated approximation errors. However, the modifications involve other approximations and lead to an SDP beamforming problem with a computational complexity that is higher compared to the proposed FIR beamforming problems based on a frequency discretization.

We also develop and analyze the DFT beamformer counterparts of the proposed FIR beamformers. These DFT beamformers are based on the narrowband beamformer of [VGL03]. The DFT beamformers differ in that the narrowband beamforming problems of one DFT beamformer contain an additional norm-constraint.

Simulation results demonstrate that the proposed beamformers are attractive alternatives to the state-of-the-art broadband adaptive beamformers. In particular, the proposed FIR beamformers provide a significantly improved performance as compared to previous FIR beamformers based on worst-case output power minimization.

5.1.1 FIR beamformers

The structure of the FIR presteered broadband beamformer with N sensors and FIR filters of length M_T is depicted in Figure 5.1.

The sensor locations are assumed to satisfy the Nyquist spatial sampling criterion, i.e., the distances between the sensors are small enough to avoid spatial aliasing effects at all frequencies. For a ULA, this corresponds to the condition that the sensors are placed at a distance less than or equal to $c/(2f_u)$, where c is the propagation speed of the waveform, $f_u = f_c + B_s/2$ is the maximum passband frequency, f_c is the carrier frequency, and B_s is the signal bandwidth [vT02]. The passband waveform received by the nth sensor is delayed by the presteering delay $T_{p,n}$. After the presteering stage, the

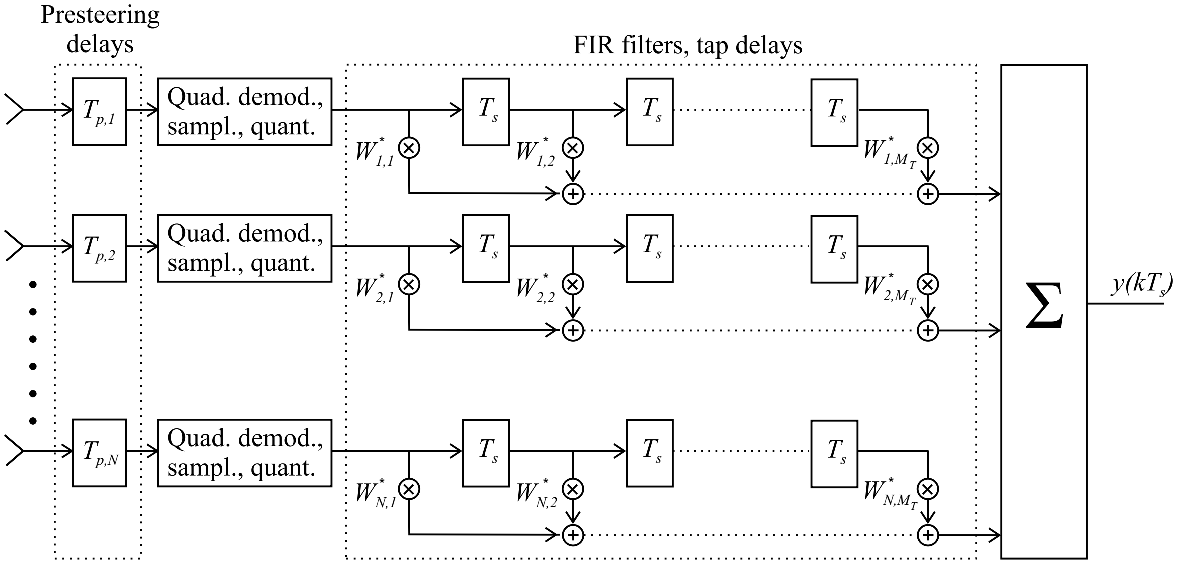


Figure 5.1. FIR presteered broadband beamformer.

passband signals are converted to baseband by means of quadrature demodulators. The baseband signals are then sampled, quantized, and filtered by FIR filters of length M_T. The tap delays T_s of the FIR filters are chosen less than or equal to $1/B_s$ to avoid aliasing in the time domain.

The complex $N \times M_T$ weight matrix $\boldsymbol{W}^*$ contains the FIR filter coefficients, where $W^*_{n,m}$ is the mth filter coefficient of the nth sensor, and $(\cdot)^*$ denotes the complex conjugate. Stacking the columns of $\boldsymbol{W}$ on top of each other leads to the $M_T N \times 1$ weight vector

$$\underline{\boldsymbol{w}} = \operatorname{vec}\{\boldsymbol{W}\}. \tag{5.1}$$

The underscore is used to distinguish the $M_T N \times 1$ weight vector $\underline{\boldsymbol{w}}$ of the FIR beamformer from the $N \times 1$ weight vector $\boldsymbol{w}$ used in the narrowband case. Similarly, the complex $M_T N \times 1$ beamformer snapshot vector at the kth time instant is given by

$$\underline{\boldsymbol{x}}(kT_s) = \begin{bmatrix} \boldsymbol{x}(kT_s) \\ \boldsymbol{x}((k-1)T_s) \\ \vdots \\ \boldsymbol{x}((k-M_T+1)T_s) \end{bmatrix}. \tag{5.2}$$

Note that FIR beamformers are often implemented without a down-conversion to baseband. The down-conversion to baseband results in complex beamformer weight vectors, and complex arithmetics require about a factor of four times as many floating point operations than real arithmetics. However, the down-conversion to baseband allows reducing the sampling rate and the FIR filter length due to the Nyquist sampling criterion. Since the computational complexities of the proposed FIR beamformers grow asymptotically as M_T^3 or faster, it can be expected that the down-conversion to baseband leads to a reduced computational complexity for the proposed FIR beamformers.

The received snapshot vector $\underline{\boldsymbol{x}}(kT_s)$ is given by

$$\underline{\boldsymbol{x}}(kT_s) = \sum_{l=1}^{L} \underline{\boldsymbol{x}}_l(kT_s) + \underline{\boldsymbol{n}}(kT_s), \tag{5.3}$$

where $\underline{\boldsymbol{x}}_l(kT_s)$ is the component due to the lth source, L denotes the number of sources, and $\underline{\boldsymbol{n}}(kT_s)$ is the noise vector. Without loss of generality, we assume that the source with index $l = 1$ is the source-of-interest, whereas the sources with indices $l = 2, \ldots, L$ are interferers.

The beamformer output can be expressed as

$$y(kT_s) = \underline{\boldsymbol{w}}^H \underline{\boldsymbol{x}}(kT_s), \tag{5.4}$$

and the average output power is

$$P_{\text{out}} = \underline{\boldsymbol{w}}^H \underline{\boldsymbol{R}}_{\boldsymbol{x}} \underline{\boldsymbol{w}}, \tag{5.5}$$

where $\underline{\boldsymbol{R}}_{\boldsymbol{x}} = \mathrm{E}\{\underline{\boldsymbol{x}}(kT_s)\underline{\boldsymbol{x}}^H(kT_s)\}$ is the $M_T N \times M_T N$ spatio-temporal snapshot covariance matrix.

The source and noise waveforms are modeled as stationary zero-mean uncorrelated random processes. Then,

$$\underline{\boldsymbol{R}}_{\boldsymbol{x}} = \sum_{l=1}^{L} \underline{\boldsymbol{R}}_l + \underline{\boldsymbol{R}}_{\boldsymbol{n}}, \tag{5.6}$$

where $\underline{\boldsymbol{R}}_l = \mathrm{E}\{\underline{\boldsymbol{x}}_l(kT_s)\underline{\boldsymbol{x}}_l^H(kT_s)\}$ is the component due to the lth source, and $\underline{\boldsymbol{R}}_{\boldsymbol{n}}$ is the noise component. The component due to the lth source can be expressed as [vT02]

$$\underline{\boldsymbol{R}}_l = \int_{-1/(2T_s)}^{1/(2T_s)} S_l(f) \underline{\boldsymbol{a}}_l(f) \underline{\boldsymbol{a}}_l^H(f) \mathrm{d}f, \tag{5.7}$$

where $S_l(f)$ is the power spectral density of the baseband waveform of the lth source,

$$\underline{\boldsymbol{a}}_l(f) = \boldsymbol{d}(f) \otimes (\boldsymbol{a}_l(f + f_c) \odot \boldsymbol{t}(f + f_c)) \tag{5.8}$$

is the frequency response vector comprising the effects of the channel, the presteering stage, and the tap delays, f is the baseband frequency, $\boldsymbol{a}_l(f+f_c)$ is the $N \times 1$ steering vector,

$$\boldsymbol{t}(f+f_c) = \left[e^{-j2\pi(f+f_c)T_{p,1}}, \ldots, e^{-j2\pi(f+f_c)T_{p,N}}\right]^T \tag{5.9}$$

is the $N \times 1$ frequency response vector of the presteering stage,

$$\boldsymbol{d}(f) = \left[1, e^{-j2\pi f T_s}, \ldots, e^{-j2\pi f(M_T-1)T_s}\right]^T, \tag{5.10}$$

$\otimes$ is the Kronecker matrix product, and $\odot$ is the Schur-Hadamard element-wise matrix product. The noise baseband waveforms are assumed to be temporally white, so

$$\underline{\boldsymbol{R}}_{\boldsymbol{n}} = \sigma \boldsymbol{I}_{M_T N}, \tag{5.11}$$

where σ is the sensor noise power.

The spatio-temporal snapshot covariance matrix can be estimated as

$$\hat{\underline{\boldsymbol{R}}}_{\boldsymbol{x}} = \frac{1}{K - M_T + 1} \sum_{k=M_T}^{K} \underline{\boldsymbol{x}}(kT_s)\underline{\boldsymbol{x}}^H(kT_s), \tag{5.12}$$

where K is the number of snapshot vectors $\boldsymbol{x}(kT_s)$.

The beamformer frequency response for the desired signal is given by

$$H(f, \boldsymbol{a}_1(f+f_c)) = \boldsymbol{a}_1^T(f+f_c)(\boldsymbol{t}(f+f_c) \odot (\boldsymbol{W}^* \boldsymbol{d}(f))), \tag{5.13}$$

where the components of the $N \times 1$ vector $\boldsymbol{W}^*\boldsymbol{d}(f)$ are the frequency responses of the FIR filters.

In practice, the steering vector for the desired signal is not known exactly due to, for example, signal wavefront distortions, signal pointing errors, or array calibration errors. Let

$$\hat{\boldsymbol{a}}_1(f+f_c) = \boldsymbol{a}_1(f+f_c) + \boldsymbol{\delta}_1(f+f_c) \tag{5.14}$$

denote the presumed (estimated) steering vector of the desired signal, where $\boldsymbol{\delta}_1(f+f_c)$ is the mismatch vector. Consequently, the beamformer frequency response for the desired signal can be expressed as

$$\begin{aligned} H(f, \boldsymbol{a}_1(f+f_c)) = {} & \hat{\boldsymbol{a}}_1^T(f+f_c)(\boldsymbol{t}(f+f_c) \odot (\boldsymbol{W}^* \boldsymbol{d}(f))) \\ & - \boldsymbol{\delta}_1^T(f+f_c)(\boldsymbol{t}(f+f_c) \odot (\boldsymbol{W}^* \boldsymbol{d}(f))). \end{aligned} \tag{5.15}$$

The estimate of the steering vector of the desired signal is modeled as [vT02]

$$\hat{\boldsymbol{a}}_1(f+f_c) = \left[e^{-j2\pi(f+f_c)\hat{\tau}_1}, \ldots, e^{-j2\pi(f+f_c)\hat{\tau}_N}\right]^T, \tag{5.16}$$

where $\hat{\tau}_n$ is the estimated propagation delay to the nth sensor relative to the propagation delay to some reference point. For example, in the case of far-field sources

$$\hat{\tau}_n = -\frac{p_{x_n}}{c}\cos(\hat{\theta}_1)\cos(\hat{\varphi}_1) - \frac{p_{y_n}}{c}\cos(\hat{\theta}_1)\sin(\hat{\varphi}_1) - \frac{p_{z_n}}{c}\sin(\hat{\theta}_1), \quad n = 1, \ldots, N, \tag{5.17}$$

where $\hat{\varphi}_1$ and $\hat{\theta}_1$ are the estimated azimuth and elevation angles of the desired signal, and $(p_{x_n}, p_{y_n}, p_{z_n})$ are the coordinates of the nth sensor. Figure 2.1 depicts the coordinate system.

The presteering delays are chosen to time-align the desired signal components in the different sensors. Hence,

$$T_{p,n} = \bar{T}_p - \hat{\tau}_n, \quad n = 1, \ldots, N, \tag{5.18}$$

where $\bar{T}_p$ is introduced to ensure causality. Substituting (5.18) in (5.9), we obtain

$$\hat{\boldsymbol{a}}_1(f + f_c) \odot \boldsymbol{t}(f + f_c) = e^{-j2\pi(f+f_c)\bar{T}_p}\mathbf{1}_N, \tag{5.19}$$

where $\mathbf{1}_N$ is an $N \times 1$ vector of ones. Then, the beamformer frequency response for the desired signal can be written as

$$\begin{aligned} H(f, \boldsymbol{a}_1(f + f_c)) = e^{-j2\pi(f+f_c)\bar{T}_p}\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}(f) \\ - \boldsymbol{\delta}_1^T(f + f_c)(\boldsymbol{t}(f + f_c) \odot (\boldsymbol{W}^*\boldsymbol{d}(f))). \end{aligned} \tag{5.20}$$

The *nominal* beamformer frequency response for the signal-of-interest is

$$H(f, \hat{\boldsymbol{a}}_1(f + f_c)) = e^{-j2\pi(f+f_c)\bar{T}_p}\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}(f). \tag{5.21}$$

Similar to (4.16), we define the FIR beamformer sensitivity to signal steering vector estimation errors at the frequency f as

$$\underline{T}_{\mathrm{se}}(f) \triangleq \frac{\|\boldsymbol{W}^*\boldsymbol{d}(f)\|_2^2}{|\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}(f)|^2}, \tag{5.22}$$

where $\|\cdot\|_2$ denotes the Euclidean vector norm. The smaller $\underline{T}_{\mathrm{se}}(f)$, the more robust is the beamformer against signal steering vector estimation errors at the frequency f. The Cauchy-Schwarz inequality implies that

$$\underline{T}_{\mathrm{se}}(f) \geq \frac{1}{N}. \tag{5.23}$$

5.1.1.1 Broadband MVDR beamformer

The MVDR beamformer minimizes the output power subject to a distortionless constraint for the desired signal. Suppose the desired signal components in the different sensors are identical after the presteering stage. Then, the distortionless constraint can be formulated as

$$\mathbf{1}_N^T \boldsymbol{W}^* = \boldsymbol{e}_{M_0}^T, \tag{5.24}$$

where $\boldsymbol{e}_{M_0}$ is a column-vector of conformable dimension containing all zeroes except for the M_0th entry, which is equal to one. The filter-and-sum stage then leads to a delay of $(M_0 - 1)T_s$ for the signal-of-interest. The corresponding desired frequency response for the signal-of-interest is given by

$$H_{1,\mathrm{des}}(f) = e^{-j2\pi((f+f_c)\bar{T}_p + f(M_0-1)T_s)}. \tag{5.25}$$

Unless specified otherwise, we assume in the following that M_T is odd, and that

$$M_0 = \frac{M_T + 1}{2}. \tag{5.26}$$

Then, the desired beamformer output is the signal-of-interest component in the central FIR filter section. The interferer and noise components in this filter section can be suppressed using both preceding and succeeding samples.

The constraint (5.24) can be reformulated as

$$(\boldsymbol{I}_{M_T} \otimes \mathbf{1}_N^T)\underline{\boldsymbol{w}} = \boldsymbol{e}_{M_0}, \tag{5.27}$$

where we have used that [Neu69]

$$\mathrm{vec}\{\boldsymbol{M}_1\boldsymbol{M}_2\boldsymbol{M}_3\} = (\boldsymbol{M}_3^T \otimes \boldsymbol{M}_1)\,\mathrm{vec}\{\boldsymbol{M}_2\} \tag{5.28}$$

for any matrices $\boldsymbol{M}_1$, $\boldsymbol{M}_2$, and $\boldsymbol{M}_3$ of conformable dimensions. Hence, the broadband MVDR beamforming problem can be written as

$$\min_{\underline{\boldsymbol{w}}} \underline{\boldsymbol{w}}^H \underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}} \underline{\boldsymbol{w}} \quad \text{s.t.} \quad \underline{\boldsymbol{C}}_{\mathrm{MVDR}}^T \underline{\boldsymbol{w}} = \boldsymbol{e}_{M_0}, \tag{5.29}$$

where $\underline{\boldsymbol{C}}_{\mathrm{MVDR}} = \boldsymbol{I}_{M_T} \otimes \mathbf{1}_N$ is the broadband MVDR constraint matrix. If $\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}$ is nonsingular, then the solution of (5.29) is given by

$$\underline{\boldsymbol{w}}_{\mathrm{MVDR}} = \underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}^{-1} \underline{\boldsymbol{C}}_{\mathrm{MVDR}} \left(\underline{\boldsymbol{C}}_{\mathrm{MVDR}}^T \underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}^{-1} \underline{\boldsymbol{C}}_{\mathrm{MVDR}} \right)^{-1} \boldsymbol{e}_{M_0}. \tag{5.30}$$

If the snapshot vectors that are used in (5.12) contain the desired signal component, then errors in the estimated signal steering vectors, presteering delay quantization errors, as well as finite sample errors can lead to a severe degradation of the broadband MVDR beamformer performance due to signal self-nulling. Multiple techniques have been proposed to improve the robustness of the broadband MVDR beamformer against such model errors. Conceptually, these techniques are often similar to the techniques that are used to improve the robustness of the narrowband MVDR beamformer.

5.1.1.2 Derivative mainlobe constraints

A popular technique to stabilize the beamformer power response towards the desired signal is to force its derivative (and possibly also higher-order derivatives) to zero. The derivative is taken with respect to uncertain system parameters such as the signal DOA [EC83, BG86, TCL93, TLT95, TCL95] or the presteering delays [ZT02].

For example, the first-order directional derivative constraints ensure that the partial derivatives of the nominal beamformer power response with respect to the azimuth and elevation angles are equal to zero for all frequencies, i.e.,

$$\frac{\partial|H(f,\hat{\boldsymbol{a}}_1(f+f_c))|^2}{\partial\hat{\theta}_1}=\frac{\partial|H(f,\hat{\boldsymbol{a}}_1(f+f_c))|^2}{\partial\hat{\varphi}_1}=0\quad\forall f\in[-B_s/2,B_s/2]\,. \tag{5.31}$$

Similar to the derivation in [EC83] for beamformers without down-conversion to baseband, it can be shown that necessary and sufficient conditions for (5.31) are

$$\begin{aligned}\mathbf{1}_N^T\boldsymbol{T}_\theta\boldsymbol{W}^*\boldsymbol{e}_{M_0+k}&=\mathbf{1}_N^T\boldsymbol{T}_\theta\boldsymbol{W}\boldsymbol{e}_{M_0-k},\quad k=0,\dots,M_0-1,\\ \mathbf{1}_N^T\boldsymbol{T}_\varphi\boldsymbol{W}^*\boldsymbol{e}_{M_0+k}&=\mathbf{1}_N^T\boldsymbol{T}_\varphi\boldsymbol{W}\boldsymbol{e}_{M_0-k},\quad k=0,\dots,M_0-1,\end{aligned} \tag{5.32}$$

where

$$\boldsymbol{T}_\theta=\operatorname{diag}\left\{\frac{\partial\hat{\tau}_1}{\partial\hat{\theta}_1},\dots,\frac{\partial\hat{\tau}_N}{\partial\hat{\theta}_1}\right\} \tag{5.33}$$

$$\boldsymbol{T}_\varphi=\operatorname{diag}\left\{\frac{\partial\hat{\tau}_1}{\partial\hat{\varphi}_1},\dots,\frac{\partial\hat{\tau}_N}{\partial\hat{\varphi}_1}\right\}. \tag{5.34}$$

Note that (5.32) contains $2M_T$ real-valued linear constraints on the components of $\boldsymbol{W}$. These constraints lead to robustness against estimation errors in the signal azimuth and elevation DOAs. Hence, the directional derivative constraints provide robustness against structured phase-errors in the estimated signal steering vectors.

Errors in the presteering delays are comparable to arbitrary-type phase errors in the estimated signal steering vectors. Hence, robustness against arbitrary-type phase errors in the estimated signal steering vectors can be achieved by means of the first-order presteering derivative constraints [ZT02]

$$\frac{\partial |H(f, \hat{\boldsymbol{a}}_1(f+f_c))|^2}{\partial T_{p,1}} = \ldots = \frac{\partial |H(f, \hat{\boldsymbol{a}}_1(f+f_c))|^2}{\partial T_{p,N}} = 0 \quad \forall f \in [-B_s/2, B_s/2]\,. \tag{5.35}$$

Similar to the derivation in [ZT02] for systems without down-conversion to the baseband, it can be shown that necessary and sufficient conditions for the first-order presteering derivative constraints are

$$\boldsymbol{W}^* \boldsymbol{e}_{M_0+k} = \boldsymbol{W} \boldsymbol{e}_{M_0-k}, \quad k = 0, \ldots, M_0 - 1. \tag{5.36}$$

Hence, these derivative constraints lead to linear phase constraints on each of the N FIR filters of the array processor. The constraints (5.36) contain $M_T N$ real-valued linear constraints. Therefore, these constraints reduce the degrees of freedom of $\boldsymbol{W}$ by a factor of two. Clearly, the presteering derivative constraints are more restrictive than the directional derivative constraints, but the former constraints provide a more general type of robustness.

Both the directional and the presteering derivative constraints do not provide robustness against magnitude errors in the components of the signal steering vectors, and it is hardly possible to control the amount of robustness.

5.1.1.3 Norm constraints

Another popular technique to improve the robustness of the broadband MVDR beamformer against signal self-nulling is to limit the norm of the beamformer weight vector [CZO87,FM94]. The norm-bounded broadband MVDR beamformer can be formulated as

$$\begin{aligned} \min_{\underline{\boldsymbol{w}}} \quad & \underline{\boldsymbol{w}}^H \hat{\underline{\boldsymbol{R}}}_{\boldsymbol{x}} \underline{\boldsymbol{w}} \\ \text{s.t.} \quad & \underline{\boldsymbol{C}}_{\text{MVDR}}^T \underline{\boldsymbol{w}} = \boldsymbol{e}_{M_0} \\ & \underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}} \leq \alpha_w, \end{aligned} \tag{5.37}$$

where α_w is a user-defined parameter.

The distortionless constraint in the second line of (5.37) implies that

$$\boldsymbol{1}_N^T \boldsymbol{W}^* \boldsymbol{d}(f) = e^{-j2\pi f(M_0-1)T_s} \quad \forall\, f \in \left[-\frac{1}{2T_s}, \frac{1}{2T_s}\right]. \tag{5.38}$$

The sensitivity of the norm-bounded broadband MVDR beamformer is therefore given by

$$\underline{T}_{\text{se}}(f) = \|\boldsymbol{W}^* \boldsymbol{d}(f)\|_2^2. \tag{5.39}$$

The norm-constraint in the last line of (5.37) leads to an upper bound on the sensitivity. Since

$$\boldsymbol{U}_{\boldsymbol{d}(f)} = \frac{1}{\sqrt{M_T}} \left[\boldsymbol{d}(f), \boldsymbol{d}\left(f + \frac{1}{M_T T_s}\right), \ldots, \boldsymbol{d}\left(f + \frac{M_T - 1}{M_T T_s}\right) \right] \tag{5.40}$$

is an $M_T \times M_T$ unitary matrix, we have

$$\|\underline{\boldsymbol{w}}\|_2^2 = \left\|\boldsymbol{W}^* \boldsymbol{U}_{\boldsymbol{d}(f)}\right\|_F^2 = \frac{1}{M_T} \sum_{m=0}^{M_T - 1} \left\| \boldsymbol{W}^* \boldsymbol{d}\left(f + \frac{m}{M_T T_s}\right) \right\|_2^2, \tag{5.41}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Applying the Cauchy-Schwarz inequality to (5.38) yields

$$\|\boldsymbol{W}^* \boldsymbol{d}(f)\|_2^2 \geq \frac{1}{N} \quad \forall\, f \in \left[-\frac{1}{2T_s}, \frac{1}{2T_s}\right]. \tag{5.42}$$

From (5.41) and (5.42), we obtain

$$\|\underline{\boldsymbol{w}}\|_2^2 \geq \frac{1}{M_T} \|\boldsymbol{W}^* \boldsymbol{d}(f)\|_2^2 + \frac{M_T - 1}{M_T N}. \tag{5.43}$$

The sensitivity of the norm-bounded broadband MVDR beamformer is therefore bounded by

$$\frac{1}{N} \leq \underline{T}_{\mathrm{se}}(f) \leq \frac{1}{N} + \left(\alpha_w - \frac{1}{N}\right) M_T, \tag{5.44}$$

where the upper bound follows from (5.43) and the norm constraint in (5.37). The lower bound has been stated in (5.23).

Due to (5.44), (5.37) is infeasible if $\alpha_w < 1/N$. The quiescent weight vector [vT02]

$$\underline{\boldsymbol{w}}_q = \boldsymbol{e}_{M_0} \otimes \mathbf{1}_N / N \tag{5.45}$$

satisfies the constraints of (5.37) if $\alpha_w = 1/N$. Hence, the norm-bounded broadband MVDR beamforming problem is feasible if and only if $\alpha_w \geq 1/N$. As in the narrowband case, choosing α_w moderately (1–3 dB) above its lower bound often leads to good performance results.

The problem in (5.37) is an SOCP problem. To solve this problem, it can be exploited that the distortionless constraint is satisfied if and only if the weight vector can be expressed as

$$\underline{\boldsymbol{w}} = \underline{\boldsymbol{w}}_q + \boldsymbol{F} \underline{\boldsymbol{w}}_{\boldsymbol{F}}, \tag{5.46}$$

where $\underline{\boldsymbol{w}}_{\boldsymbol{F}}$ is an $M_T(N-1) \times 1$ vector, and $\boldsymbol{F}$ is an $M_T N \times M_T(N-1)$ matrix whose columns span the orthogonal complement of the column-space of $\underline{\boldsymbol{C}}_{\mathrm{MVDR}}$, i.e.,

$$\underline{\boldsymbol{C}}_{\mathrm{MVDR}}^H \boldsymbol{F} = \mathbf{0}. \tag{5.47}$$

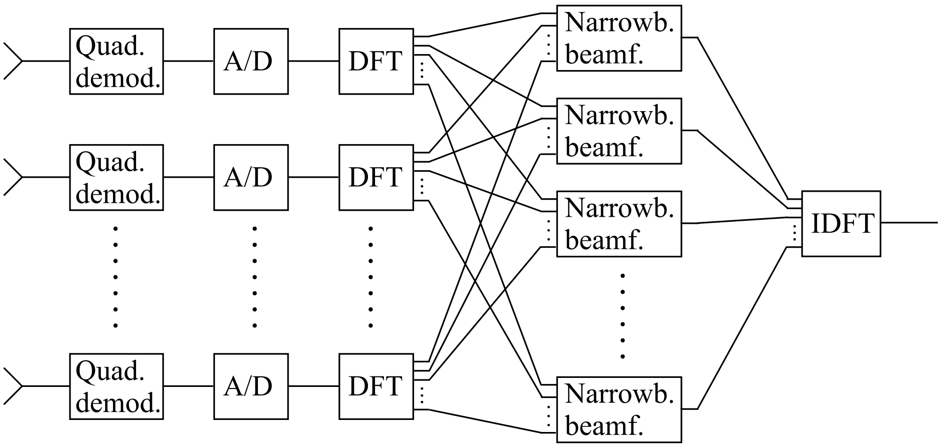


Figure 5.2. DFT beamformer.

Therefore, the optimization problem (5.37) can be written equivalently as

$$\begin{aligned} \min_{\underline{\boldsymbol{w}}_{\boldsymbol{F}}} \quad & (\underline{\boldsymbol{w}}_q^H + \underline{\boldsymbol{w}}_{\boldsymbol{F}}^H \boldsymbol{F}^H) \hat{\underline{\boldsymbol{R}}}_{\boldsymbol{x}} (\underline{\boldsymbol{w}}_q + \boldsymbol{F} \underline{\boldsymbol{w}}_{\boldsymbol{F}}) \\ \text{s.t.} \quad & \|\underline{\boldsymbol{w}}_q + \boldsymbol{F} \underline{\boldsymbol{w}}_{\boldsymbol{F}}\|_2^2 \leq \alpha_w . \end{aligned} \tag{5.48}$$

The solution of (5.48) can be obtained efficiently by a 1D search over the Lagrange parameter, similar as in [LSW03]. The computational complexity of (5.48) is $\mathcal{O}(M_T^3 N^3)$.

5.1.2 DFT beamformers

Figure 5.2 depicts the DFT beamformer structure [vT02]. The received passband waveforms are down-converted to baseband by means of quadrature demodulators. Subsequently, the baseband waveforms are sampled and quantized. Then, DFTs of length Q are applied to obtain narrowband frequency components. For DFT beamformers, we consider the uniform frequency grid

$$f_q = -\frac{1}{2T_s} + \frac{q-1}{QT_s}, \quad q = 1, \ldots, Q. \tag{5.49}$$

The frequency domain snapshot vectors for the frequency f_q are computed as

$$\boldsymbol{x}_q(kT_s) = \frac{1}{Q} \sum_{m=1}^{Q} \alpha_m \boldsymbol{x}((k-m+1)T_s)\, e^{j2\pi f_q (m-1) T_s}, \quad k = Q, \ldots, K, \tag{5.50}$$

where α_m, $m = 1, \ldots, Q$ are windowing coefficients [OS99].

The $N \times N$ covariance matrix of the frequency domain snapshot vectors for the frequency f_q can be estimated as

$$\hat{\boldsymbol{R}}_{\boldsymbol{x}_q} = \frac{1}{K-Q+1} \sum_{k=Q}^{K} \boldsymbol{x}_q(kT_s)\boldsymbol{x}_q^H(kT_s). \tag{5.51}$$

The presumed signal steering vector for the qth subband is $\hat{\boldsymbol{a}}_1(f_q + f_c)$. Based on these estimates, the beamformer weight vector for the qth subband can be computed using narrowband beamforming techniques.

In the sequel, we consider the sliding window implementation of DFT beamformers [vT02]. That means, the DFTs are computed at each sampling time using the Q most recent snapshot vectors. Then, the IDFT can be reduced to a weighted summation of the narrowband beamformer output samples. For a discussion of the similar DFT beamformer implementation based on block processing see, for example, [vT02].

Let $\boldsymbol{w}_q$ denote the optimum weight vector for the qth subband. The time-domain beamformer output samples are computed as

$$y(kT_s) = \sum_{q=1}^{Q} \beta_q \boldsymbol{w}_q^H \boldsymbol{x}_q(kT_s) e^{-j2\pi f_q (Q_0-1)T_s}, \tag{5.52}$$

where the scaling factors

$$\beta_q = \frac{1}{\boldsymbol{w}_q^H \hat{\boldsymbol{a}}_1(f_q + f_c)}, \quad q = 1, \ldots, Q \tag{5.53}$$

ensure that the nominal frequency response for the signal-of-interest is equal to the desired distortionless frequency response, and the delay of the signal-of-interest is $(Q_0 - 1)T_s$ in the error-free case. In the sequel, we choose

$$Q_0 = \left\lfloor \frac{Q+1}{2} \right\rfloor, \tag{5.54}$$

where $\lfloor \cdot \rfloor$ is the floor function, which maps a real number to the largest previous integer. Note that Q_0 has a similar impact than M_0 for FIR beamformers.

Due to (5.50), the frequency domain snapshot vectors at the kth time instant can be expressed as

$$[\boldsymbol{x}_1(kT_s), \ldots, \boldsymbol{x}_Q(kT_s)] = \underline{\underline{\boldsymbol{X}}}(kT_s)\, \boldsymbol{D}_\alpha\, \boldsymbol{Q}_{\mathrm{DFT}}, \tag{5.55}$$

where

$$\underline{\underline{\boldsymbol{X}}}(kT_s) = [\boldsymbol{x}(kT_s), \ldots, \boldsymbol{x}((k-Q+1)T_s)] \tag{5.56}$$

is the $N \times Q$ snapshot vector matrix,

$$\boldsymbol{D}_{\alpha} = \operatorname{diag}\{\alpha_1, \ldots, \alpha_Q\} \tag{5.57}$$

contains the window coefficients, and

$$\boldsymbol{Q}_{\mathrm{DFT}} = \frac{1}{Q} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{j2\pi f_1 T_s} & e^{j2\pi f_2 T_s} & \cdots & e^{j2\pi f_Q T_s} \\ \vdots & \vdots & & \vdots \\ e^{j2\pi f_1 (Q-1) T_s} & e^{j2\pi f_2 (Q-1) T_s} & \cdots & e^{j2\pi f_Q (Q-1) T_s} \end{bmatrix} \tag{5.58}$$

is the $Q \times Q$ DFT matrix. Let

$$\underline{\underline{\boldsymbol{W}}} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_Q] \tag{5.59}$$

contain the narrowband beamformer weight vectors. Then,

$$\underline{\underline{\boldsymbol{W}}}^H \, \underline{\underline{\boldsymbol{X}}}(kT_s) \, \boldsymbol{D}_{\alpha} \, \boldsymbol{Q}_{\mathrm{DFT}} \tag{5.60}$$

is a $Q \times Q$ matrix, which contains the narrowband beamformer output samples on its diagonal. Equation (5.52) can therefore be written as

$$y(kT_s) = \operatorname{tr}\left\{ \boldsymbol{D}_{\mathrm{IDFT}} \, \boldsymbol{D}_{\beta} \, \underline{\underline{\boldsymbol{W}}}^H \, \underline{\underline{\boldsymbol{X}}}(kT_s) \, \boldsymbol{D}_{\alpha} \, \boldsymbol{Q}_{\mathrm{DFT}} \right\}, \tag{5.61}$$

where

$$\begin{aligned} \boldsymbol{D}_{\beta} &= \operatorname{diag}\{\beta_1, \ldots, \beta_Q\} \\ \boldsymbol{D}_{\mathrm{IDFT}} &= \operatorname{diag}\left\{ e^{-j2\pi f_1 (Q_0 - 1) T_s}, \ldots, e^{-j2\pi f_Q (Q_0 - 1) T_s} \right\}, \end{aligned} \tag{5.62}$$

and $\operatorname{tr}\{\cdot\}$ denotes the trace operator. Hence, (5.61) can be reformulated as

$$y(kT_s) = \underline{\underline{\boldsymbol{w}}}^H \underline{\underline{\boldsymbol{x}}}(kT_s), \tag{5.63}$$

where

$$\underline{\underline{\boldsymbol{x}}}(kT_s) = \operatorname{vec}\left\{ \underline{\underline{\boldsymbol{X}}}(kT_s) \right\} \tag{5.64}$$

is the $NQ \times 1$ spatio-temporal snapshot vector, and

$$\underline{\underline{\boldsymbol{w}}}(kT_s) = \operatorname{vec}\left\{ \underline{\underline{\boldsymbol{W}}} \, \boldsymbol{D}_{\beta}^{*} \, \boldsymbol{D}_{\mathrm{IDFT}}^{*} \, \boldsymbol{Q}_{\mathrm{DFT}}^{H} \boldsymbol{D}_{\alpha} \right\} \tag{5.65}$$

is the DFT beamformer weight vector.

The computational complexity of DFT beamformers is usually lower as compared to the FIR beamformer counterparts with $M_T = Q$. For example, the complexity of the broadband MVDR beamformer (5.29) is $\mathcal{O}(M_T^3 N^3)$, whereas the complexity of the DFT beamformer using narrowband MVDR techniques is $\mathcal{O}(NQ \log Q + QN^3)$.

The results of DFT beamformers are non-optimum, since correlations between the frequency domain snapshot vectors of different subbands are not taken into account. Such correlations can be reduced by increasing the DFT length. This also provides an improved resolution capability in the frequency domain [Com88].

5.2 FIR beamformers based on worst-case output power minimization

In Sections 5.2.1 and 5.2.2, we propose two FIR beamformers based on worst-case output power minimization, which use different constraints to prevent distortions of the desired signal. In Section 5.2.3, we study the relation between the proposed FIR beamformers. In Section 5.2.4, we establish a relation between the proposed FIR beamformers and the norm-bounded broadband MVDR beamformer. The FIR beamformers proposed in Sections 5.2.1 and 5.2.2 control the frequency response towards the desired signal only for a finite set of frequencies. In Section 5.2.5, we develop a modified version of the first proposed FIR beamformer, which avoids such a frequency discretization and associated errors.

5.2.1 First proposed FIR beamformer

The beamformer frequency response for the signal-of-interest should approximately be equal to $H_{1,\mathrm{des}}(f)$. Since $H_{1,\mathrm{des}}(f)$ is unit modulus,

$$H(f, \boldsymbol{a}_1(f+f_c))H^*_{1,\mathrm{des}}(f) \tag{5.66}$$

should approximately be equal to one.

The signal steering vector $\boldsymbol{a}_1(f+f_c)$ is not known exactly, but an upper bound on the norm of the error vector $\boldsymbol{\delta}_1(f+f_c)$ can be estimated based on assumptions on the array calibration errors, signal pointing errors, sensor imperfections, etc. We denote this upper bound by $\epsilon_1(f+f_c)$ and assume in the following that

$$\epsilon_1(f+f_c) < \|\hat{\boldsymbol{a}}_1(f+f_c)\|_2. \tag{5.67}$$

To prevent the suppression of the desired signal, we use the constraints

$$\begin{aligned} &\Re\left\{H(f, \hat{\boldsymbol{a}}_1(f+f_c)+\tilde{\boldsymbol{\delta}})H^*_{1,\mathrm{des}}(f)\right\} \geq \gamma(f) \\ &\forall \left\|\tilde{\boldsymbol{\delta}}\right\|_2 \leq \epsilon_1(f+f_c),\ f \in [-B_s/2, B_s/2], \end{aligned} \tag{5.68}$$

where $\gamma(f)$ is a strictly positive user-defined lower bound. Note that the narrowband beamformer based on worst-case output power minimization (4.31) enforces a similar constraint, where the lower bound has been set to one. In the case of the narrowband beamformer (4.31), the OSINR performance does not depend on the choice of the lower bound. In contrast, we will see in the following that the performance of the proposed beamformer depends on $\gamma(f)$, and that $\gamma(f) = 1$ is not necessarily the preferred choice.

The beamformer frequency response for the steering vector $\hat{\boldsymbol{a}}_1(f+f_c)+\tilde{\boldsymbol{\delta}}$ can be written as

$$\begin{aligned} H(f, \hat{\boldsymbol{a}}_1(f+f_c)+\tilde{\boldsymbol{\delta}}) = {} & e^{-j2\pi(f+f_c)\bar{T}_p}\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}(f) \\ & + \tilde{\boldsymbol{\delta}}^T\left(\boldsymbol{t}(f+f_c)\odot(\boldsymbol{W}^*\boldsymbol{d}(f))\right), \end{aligned} \tag{5.69}$$

which follows directly from (5.13), and (5.19). Hence, the constraints in (5.68) can be written as

$$\begin{aligned} & \Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f) + \tilde{\boldsymbol{\delta}}^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f)\right\} \geq \gamma(f) \\ & \forall\ \|\tilde{\boldsymbol{\delta}}\|_2 \leq \epsilon_1(f+f_c),\ f\in[-B_s/2, B_s/2]\,, \end{aligned} \tag{5.70}$$

where the $M_T \times 1$ vector

$$\boldsymbol{d}_{\mathrm{cs}}(f) = \boldsymbol{d}(f)\, e^{j2\pi f(M_0-1)T_s} \tag{5.71}$$

is conjugate symmetric.

The constraints in (5.70) are satisfied if and only if they are satisfied for the worst-case mismatch vectors, i.e., if and only if

$$\begin{aligned} \min_{\tilde{\boldsymbol{\delta}}} \quad & \Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f) + \tilde{\boldsymbol{\delta}}^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f)\right\} \\ \text{s.t.} \quad & \|\tilde{\boldsymbol{\delta}}\|_2 \leq \epsilon_1(f+f_c) \end{aligned} \tag{5.72}$$

is greater than or equal to $\gamma(f)$. The solution of (5.72) is given by

$$\Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f)\right\} - \epsilon_1(f+f_c)\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f)\|_2, \tag{5.73}$$

which is attained for the worst-case mismatch vector

$$\boldsymbol{\delta}_{\mathrm{wc}}(f+f_c) = -\frac{\boldsymbol{W}\boldsymbol{d}_{\mathrm{cs}}^*(f)}{\|\boldsymbol{W}\boldsymbol{d}_{\mathrm{cs}}^*(f)\|_2}\,\epsilon_1(f+f_c). \tag{5.74}$$

This leads to the beamforming problem

$$\begin{aligned} \min_{\boldsymbol{W}} \quad & \mathrm{vec}^H\{\boldsymbol{W}\}\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\mathrm{vec}\{\boldsymbol{W}\} \\ \text{s.t.} \quad & \Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f)\right\} - \epsilon_1(f+f_c)\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f)\|_2 \geq \gamma(f) \\ & \forall\ f\in[-B_s/2, B_s/2]\,. \end{aligned} \tag{5.75}$$

The beamformer (5.75) minimizes the worst-case output power. As in the case of the narrowband beamformer based on worst-case output power minimization (4.26), this does not imply that the beamformer (5.75) optimizes any meaningful worst-case performance for the presumed set of signal steering vectors.

The problem (5.75) contains an infinite number of convex constraints. To approximately solve this problem, we sample the frequency range, and enforce the constraints for the frequencies $f_1, \dots, f_{N_f} \in [-B_s/2, B_s/2]$. Hence, we solve

$$\begin{aligned} \min_{\boldsymbol{W}} \quad & \mathrm{vec}^H\{\boldsymbol{W}\}\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\mathrm{vec}\{\boldsymbol{W}\} \\ \text{s.t.} \quad & \Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\right\} - \epsilon_1(f_k+f_c)\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2 \geq \gamma(f_k), \\ & k = 1, \dots, N_f. \end{aligned} \tag{5.76}$$

The problem (5.76) is a SOCP problem, which can be solved, for example, by means of the logarithmic barrier method [BV04]. It is shown in Appendix A.7 that the barrier function, its gradient, and its Hessian can be computed with the complexity $\mathcal{O}(M_T^2N^2 + M_TN^2N_f)$. The complexity for the exact computation of a Newton step is $\mathcal{O}((M_TN + N_f)^3)$ due to the inversion of the Hessian matrix. The number of Newton steps required to solve (5.76) weakly depends on the problem size, and is typically smaller than 100. A standard filter design rule-of-thumb is to choose $N_f = 10M_T$ equispaced frequency samples over the interval $[-B_s/2, B_s/2]$, see [Che82, WBV97] and references therein. Consequently, the order of the computational complexity of (5.76) is comparable to that of the norm-bounded broadband MVDR beamformer.

5.2.1.1 Robustness against signal self-nulling

The constraints of the beamformer (5.76) imply that

$$\underline{T}_{\mathrm{se}}(f_k) \leq \frac{\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2^2}{\left(\gamma(f_k) + \epsilon_1(f_k+f_c)\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2\right)^2}. \tag{5.77}$$

Hence, the sensitivity of the first proposed FIR beamformer is bounded by

$$\frac{1}{N} \leq \underline{T}_{\mathrm{se}}(f_k) \leq \frac{1}{\epsilon_1^2(f_k+f_c)}, \tag{5.78}$$

with the lower bound already stated in (5.23).

Moreover, (5.20) yields

$$|H(f_k, \boldsymbol{a}_1(f_k+f_c))| \geq \Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\right\} - \|\boldsymbol{\delta}_1(f_k+f_c)\|_2\,\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2. \tag{5.79}$$

The constraints in (5.76) therefore imply that

$$|H(f_k, \boldsymbol{a}_1(f_k+f_c))| \geq \gamma(f_k) + \left(\epsilon_1(f_k+f_c) - \|\boldsymbol{\delta}_1(f_k+f_c)\|_2\right)\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2. \tag{5.80}$$

The minimization of the beamformer output power leads to small values of $|H(f_k, \boldsymbol{a}_1(f_k + f_c))|$. The larger the power spectral density of the desired signal at the frequency f_k, the stronger is the incentive for a small value of $|H(f_k, \boldsymbol{a}_1(f_k + f_c))|$. Due to (5.80), this results in small values of $\|\boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k)\|_2$ if

$$\|\boldsymbol{\delta}_1(f + f_c)\|_2 < \epsilon_1(f + f_c). \tag{5.81}$$

The upper bound on the sensitivity in (5.77) is a strictly increasing function of $\|\boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k)\|_2$. Consequently, the constraints of the beamformer (5.76) lead to an incentive for a low sensitivity to signal steering vector estimation errors if (5.81) holds. The strength of this incentive increases with the power of the desired signal. A similar signal-dependent incentive for a low sensitivity does not exist for the norm-bounded broadband MVDR beamformer.

The magnitude of the nominal frequency response for the signal-of-interest may be large for the optimum weight matrix of (5.76). If we normalize this weight matrix so that the nominal beamformer frequency response is equal to the desired distortionless frequency response, then the beamformer gain may be small for the worst-case mismatch vector. It is therefore recommended to choose $\epsilon_1(f_k + f_c)$ significantly larger than the expected norms of the signal steering vector estimation errors to avoid signal self-nulling.

5.2.1.2 Geometric interpretation

The set of beamformer frequency responses corresponding to the set of presumed signal steering vectors for the frequency f_k is

$$\mathcal{S}_{\epsilon_1}(f_k) = \left\{ H(f_k, \hat{\boldsymbol{a}}_1(f_k + f_c) + \tilde{\boldsymbol{\delta}}) \,\middle|\, \left\|\tilde{\boldsymbol{\delta}}\right\|_2 \leq \epsilon_1(f_k + f_c) \right\}. \tag{5.82}$$

This set is a sphere in the complex plane with the radius $\epsilon_1(f_k + f_c)\|\boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k)\|_2$. The center of this sphere is the nominal beamformer frequency response. The constraints in (5.76) ensure that $\mathcal{S}_{\epsilon_1}(f_k)$ lies within the halfplane

$$\mathcal{S}_{\text{hp}}(f_k) = \left\{ H | \Re \left\{ H H_{1,\text{des}}^*(f_k) \right\} \geq \gamma(f_k) \right\}. \tag{5.83}$$

This halfplane is bounded by the tangent to the circle with radius $\gamma(f_k)$

$$\mathcal{S}_{\tan}(f_k) = \left\{ H | \Re \left\{ H H_{1,\text{des}}^*(f_k) \right\} = \gamma(f_k) \right\}. \tag{5.84}$$

Figure 5.3 illustrates $\mathcal{S}_{\epsilon_1}(f_k)$. If (5.81) is satisfied, then $H(f_k, \boldsymbol{a}_1(f_k + f_c))$ lies within $\mathcal{S}_{\epsilon_1}(f_k)$. It can be expected that the minimization of the beamformer output power leads to small values of $|H(f_k, \boldsymbol{a}_1(f_k + f_c))|$. Hence, for the optimum weight matrix of (5.76), $\mathcal{S}_{\epsilon_1}(f_k)$ lies close to $\gamma(f_k) H_{1,\text{des}}(f_k)$, and the phase of $H(f_k, \boldsymbol{a}_1(f_k + f_c))$ is approximately equal to the phase of $H_{1,\text{des}}(f_k)$.

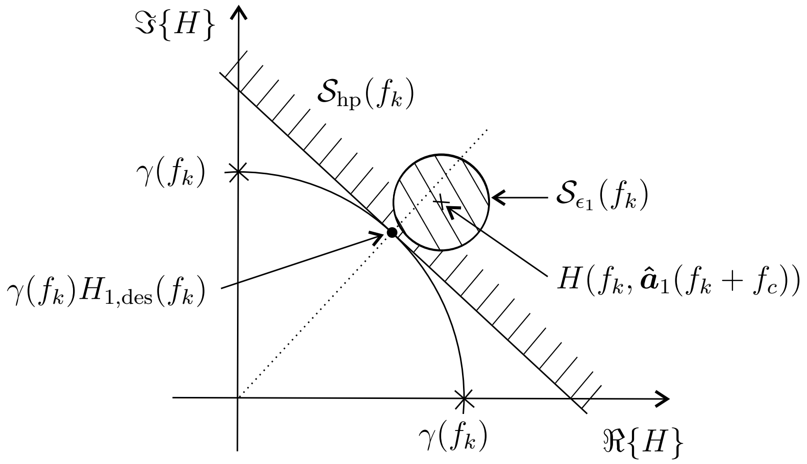


Figure 5.3. Geometric interpretation of the constraints of (5.76).

5.2.1.3 Presteering delay quantization errors

Presteering delay quantization errors may be treated as additional phase errors in the estimates of the signal steering vectors. Hence, larger values of $\epsilon_1(f_k + f_c)$ may be chosen, taking into account the additional phase errors. However, this approach is non-optimum if the presteering delay quantization errors are known.

In the case of presteering delay quantization errors, (5.19) no longer holds. Using the general vector $\hat{\boldsymbol{a}}_1(f+f_c) \odot \boldsymbol{t}(f+f_c)$ instead of $e^{-j2\pi(f+f_c)\bar{T}_p}\mathbf{1}_N$ leads to the beamformer formulation

$$\begin{aligned} \min_{\boldsymbol{W}} \quad & \mathrm{vec}^H\{\boldsymbol{W}\}\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\mathrm{vec}\{\boldsymbol{W}\} \\ \text{s.t.} \quad & \Re\left\{(\hat{\boldsymbol{a}}_1(f_k+f_c) \odot \boldsymbol{t}(f_k+f_c))^T \boldsymbol{W}^* \boldsymbol{d}(f_k) H^*_{1,\mathrm{des}}(f_k)\right\} \\ & - \epsilon_1(f_k+f_c)\|\boldsymbol{W}^*\boldsymbol{d}(f_k)\|_2 \geq \gamma(f_k), \quad k = 1, \ldots, N_f, \end{aligned} \tag{5.85}$$

which is a SOCP problem. It can be shown that the order of the computational complexity of (5.85) is equal to that of (5.76).

The beamformer (5.85) can be used even if there are no presteering delays. However, without a presteering stage, substantial delays can occur between the desired signal components in the different sensors, in particular for arrays with large apertures. These delays can be compensated by the FIR filter stage if the length of the FIR filters is sufficiently large. Since long FIR filters lead to high computational complexities, presteering delays are typically used in practice.

5.2.2 Second proposed FIR beamformer

The second proposed FIR beamformer directly limits the deviation of the beamformer frequency response from its desired distortionless value as

$$\begin{aligned} &\left|H(f_k, \hat{\boldsymbol{a}}_1(f_k + f_c) + \tilde{\boldsymbol{\delta}}) - H_{1,\text{des}}(f_k)\right| \le \zeta(f_k) \\ &\forall \left\|\tilde{\boldsymbol{\delta}}\right\|_2 \le \epsilon_1(f_k + f_c),\ k = 1, \dots, N_f, \end{aligned} \tag{5.86}$$

where $\zeta(f_k)$ is the maximum allowed deviation for the frequency f_k. Clearly, small values of $\zeta(f_k)$ lead to small distortions of the desired signal, but also limit the capability to suppress interferers and noise.

Substituting (5.69) in (5.86) and simplifying the resulting constraints yields

$$\begin{aligned} &\left|\mathbf{1}_N^T \boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k) + \tilde{\boldsymbol{\delta}}^T \boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k) - 1\right| \le \zeta(f_k) \\ &\forall \left\|\tilde{\boldsymbol{\delta}}\right\|_2 \le \epsilon_1(f_k + f_c),\ k = 1, \dots, N_f. \end{aligned} \tag{5.87}$$

The worst-case deviation from the desired distortionless frequency response is given by

$$\begin{aligned} \max_{\tilde{\boldsymbol{\delta}}} \quad & \left|\mathbf{1}_N^T \boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k) + \tilde{\boldsymbol{\delta}}^T \boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k) - 1\right| \\ \text{s.t.} \quad & \left\|\tilde{\boldsymbol{\delta}}\right\|_2 \le \epsilon_1(f_k + f_c). \end{aligned} \tag{5.88}$$

The solution of (5.88) is

$$\left|\mathbf{1}_N^T \boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k) - 1\right| + \epsilon_1(f_k + f_c)\|\boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k)\|_2. \tag{5.89}$$

Hence, the second proposed FIR beamformer can be formulated as

$$\begin{aligned} \min_{\boldsymbol{W}} \quad & \text{vec}^H\{\boldsymbol{W}\} \underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}} \text{vec}\{\boldsymbol{W}\} \\ \text{s.t.} \quad & \left|\mathbf{1}_N^T \boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k) - 1\right| + \epsilon_1(f_k + f_c)\|\boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k)\|_2 \le \zeta(f_k), \quad k = 1, \dots, N_f. \end{aligned} \tag{5.90}$$

The problem (5.76) is a SOCP problem, which can be solved using, for example, the logarithmic barrier method [BV04]. In Appendix A.8, we show that the order of the complexity to compute the barrier function, its gradient, and its Hessian is the same as for the beamformer (5.76). The first and second proposed FIR beamformers therefore have comparable computational complexities.

To avoid the trivial solution, we assume in the following that $\zeta(f_k) < 1$ for all frequencies. The following proposition provides necessary and sufficient conditions for the feasibility of (5.90).

Proposition 5.1. *The problem* (5.90) *is feasible if and only if*

$$\zeta(f_k) \geq \frac{\epsilon_1(f_k + f_c)}{\sqrt{N}} \quad \forall\, k = 1, \ldots, N_f. \tag{5.91}$$

If (5.91) *is satisfied with equality for all* $k = 1, \ldots, N_f$*, then the quiescent weight matrix [vT02]*

$$\boldsymbol{W}_q = \frac{\mathbf{1}_N}{N} \boldsymbol{e}_{M_0}^T \tag{5.92}$$

is a feasible point.

Proof. See Appendix A.9. □

5.2.2.1 Robustness against signal self-nulling

The constraints in (5.90) imply that

$$\left|\mathbf{1}_N^T \boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f_k)\right| \geq 1 - \zeta(f_k) + \epsilon_1(f_k + f_c) \|\boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2, \quad k = 1, \ldots, N_f. \tag{5.93}$$

Hence, the sensitivity of the second proposed FIR beamformer is bounded by

$$\underline{T}_{\mathrm{se}}(f_k) \leq \frac{\|\boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2^2}{\left(1 - \zeta(f_k) + \epsilon_1(f_k + f_c)\|\boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2\right)^2}. \tag{5.94}$$

The upper bound in (5.94) is a strictly increasing function of $\|\boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2$. The constraints in (5.90) give

$$\|\boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2 \leq \frac{\zeta(f_k)}{\epsilon_1(f_k + f_c)}, \quad k = 1, \ldots, N_f. \tag{5.95}$$

Therefore, the sensitivity of the second proposed FIR beamformer is bounded by

$$\frac{1}{N} \leq \underline{T}_{\mathrm{se}}(f_k) \leq \frac{\zeta^2(f_k)}{\epsilon_1^2(f_k + f_c)}, \tag{5.96}$$

where the lower bound has been stated in (5.23). Hence, the constraints of the second proposed FIR beamformer limit the sensitivity to smaller values than the constraints of the first proposed FIR beamformer.

Furthermore, (5.20) gives

$$|H(f_k, \boldsymbol{a}_1(f_k + f_c))| \geq |\mathbf{1}_N^T \boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f_k)| - \|\boldsymbol{\delta}_1(f_k + f_c)\|_2 \, \|\boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2. \tag{5.97}$$

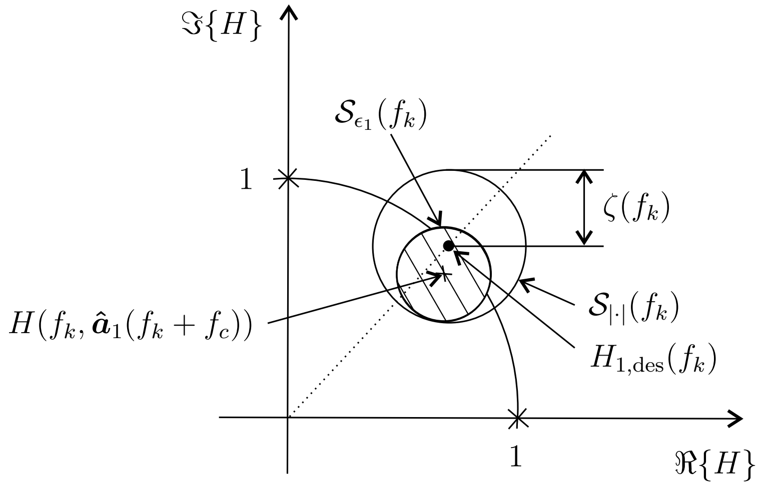


Figure 5.4. Geometric interpretation of the constraints of (5.90).

Using (5.93) and (5.97), we obtain

$$\begin{aligned} |H(f_k, \boldsymbol{a}_1(f_k + f_c))| \geq 1 - \zeta(f_k) \\ + (\epsilon_1(f_k + f_c) - \|\boldsymbol{\delta}_1(f_k + f_c)\|_2) \|\boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k)\|_2. \end{aligned} \tag{5.98}$$

The minimization of the beamformer output power leads to small values of $|H(f_k, \boldsymbol{a}_1(f_k+f_c))|$. Due to (5.98), this results in small values of $\|\boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k)\|_2$ if (5.81) holds. Since the upper bound on the beamformer sensitivity in (5.94) is a strictly increasing function of $\|\boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k)\|_2$, the minimization of the beamformer output power results in an incentive for a low sensitivity to signal steering vector estimation errors. The larger the power spectral density of the desired signal at the frequency f_k, the stronger is the incentive for a small value of $\underline{T}_{\text{se}}(f_k)$.

5.2.2.2 Geometric interpretation

The geometric interpretation of the kth constraint of the second proposed FIR beamformer (5.90) is illustrated in Figure 5.4. The set of presumed beamformer responses $\mathcal{S}_{\epsilon_1}(f_k)$ lies within the set

$$\mathcal{S}_{|\cdot|}(f_k) = \{H \mid |H - H_{1,\text{des}}(f_k)| \leq \zeta(f_k)\}, \tag{5.99}$$

which is a sphere of radius $\zeta(f_k)$ that is centered on the desired beamformer frequency response $H_{1,\text{des}}(f_k)$.

5.2.2.3 Presteering delay quantization errors

In the case of presteering delay quantization errors, similar steps as above lead to the beamforming problem

$$\begin{aligned} \min_{\boldsymbol{W}} \quad & \mathrm{vec}^H\{\boldsymbol{W}\}\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\mathrm{vec}\{\boldsymbol{W}\} \\ \text{s.t.} \quad & \left|(\hat{\boldsymbol{a}}_1(f_k+f_c)\odot \boldsymbol{t}(f_k+f_c))^T\boldsymbol{W}^*\boldsymbol{d}(f_k) - H_{1,\mathrm{des}}(f_k)\right| \\ & + \epsilon_1(f_k+f_c)\|\boldsymbol{W}^*\boldsymbol{d}(f_k)\|_2 \le \zeta(f_k), \quad k=1,\dots,N_f, \end{aligned} \tag{5.100}$$

which is a SOCP problem. It can be shown that the order of the computational complexity of (5.100) is equal to that of (5.90).

Proposition 5.1 does not apply in the case of (5.100), and it is hardly possible to formulate necessary and sufficient conditions for the feasibility of (5.100). In the case of presteering delay quantization errors, $\zeta(f_k)$ may be chosen as follows. Similar steps as in Appendix A.9 show that

$$\min_{\boldsymbol{W}} \left|(\hat{\boldsymbol{a}}_1(f_k+f_c)\odot \boldsymbol{t}(f_k+f_c))^T\boldsymbol{W}^*\boldsymbol{d}(f_k) - H_{1,\mathrm{des}}(f_k)\right| + \epsilon_1(f_k+f_c)\|\boldsymbol{W}^*\boldsymbol{d}(f_k)\|_2 \tag{5.101}$$

is attained if and only if

$$\boldsymbol{W}^*\boldsymbol{d}(f_k) = \frac{\hat{\boldsymbol{a}}_1^*(f_k+f_c)\odot \boldsymbol{t}^*(f_k+f_c)H_{1,\mathrm{des}}(f_k)}{N}. \tag{5.102}$$

In general, there is no matrix $\boldsymbol{W}$, which satisfies (5.102) for all $k=1,\dots,N_f$. However, we can compute the matrix $\boldsymbol{W}$, which minimizes the least squares deviation between the left- and right-hand side expressions of (5.102) for $k=1,\dots,N_f$. Choosing $\zeta(f_k)$ somewhat larger than the objective function in (5.101) for the least squares optimal weight matrix ensures feasibility and leads to small signal distortions.

5.2.3 Relation between the proposed FIR beamformers

In Section 5.2.2.1, we compared the sensitivity of the first and second proposed FIR beamformers. We showed that the constraints of the second proposed FIR beamformer limit the sensitivity to smaller values than the constraints of the first proposed FIR beamformer. If (5.81) holds, then the constraints of both proposed FIR beamformers lead to an incentive for a low sensitivity, which becomes stronger with increasing signal powers. In this section, we further investigate the relation between the proposed FIR beamformers. For the sake of simplicity, we assume in the following that the presteering delay quantization errors are negligible.

The constraints of the second proposed FIR beamformer (5.90) imply that

$$\Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\right\} - \epsilon_1(f_k+f_c)\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2 \geq 1-\zeta(f_k), \quad k=1,\dots,N_f, \tag{5.103}$$

where we have used that

$$\left|\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)-1\right| \geq 1-\Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\right\}. \tag{5.104}$$

The constraints of (5.90) are therefore stricter than the constraints of (5.76) if

$$1-\zeta(f_k) \geq \gamma(f_k), \quad k=1,\dots,N_f. \tag{5.105}$$

In this case, (5.76) is a relaxation of (5.90). The geometric interpretation of this result is that $\mathcal{S}_{\mathrm{hp}}(f_k)$ is a superset of $\mathcal{S}_{|\cdot|}(f_k)$ if (5.105) holds.

Moreover, it can be seen easily that adding the distortionless constraint (5.24) to the beamformer (5.76) leads to the beamforming problem

$$\begin{aligned} \min_{\underline{\boldsymbol{w}}} \quad & \underline{\boldsymbol{w}}^H\underline{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}\underline{\boldsymbol{w}} \\ \text{s.t.} \quad & 1-\epsilon_1(f_k+f_c)\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2 \geq \gamma(f_k), \quad k=1,\dots,N_f \\ & \mathbf{1}_N^T\boldsymbol{W}^* = \boldsymbol{e}_{M_0}^T. \end{aligned} \tag{5.106}$$

Adding the distortionless constraint (5.24) to the second proposed FIR beamformer (5.90) yields

$$\begin{aligned} \min_{\underline{\boldsymbol{w}}} \quad & \underline{\boldsymbol{w}}^H\underline{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}\underline{\boldsymbol{w}} \\ \text{s.t.} \quad & \epsilon_1(f_k+f_c)\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2 \leq \zeta(f_k), \quad k=1,\dots,N_f \\ & \mathbf{1}_N^T\boldsymbol{W}^* = \boldsymbol{e}_{M_0}^T. \end{aligned} \tag{5.107}$$

It can be seen easily that (5.106) and (5.107) are identical if

$$\gamma(f_k) = 1-\zeta(f_k), \quad k=1,\dots,N_f. \tag{5.108}$$

The FIR beamformers proposed in Sections 5.2.1 and 5.2.2 are therefore closely related.

Figure 5.5 depicts the geometric interpretation of the kth constraint of (5.107). Due to the distortionless constraint (5.24), $\mathcal{S}_{\epsilon_1}(f_k)$ is centered on $H_{1,\mathrm{des}}(f_k)$. The radius of $\mathcal{S}_{\epsilon_1}(f_k)$ is $\epsilon_1(f_k+f_c)\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\|_2$, which is upper bounded by $\zeta(f_k)$ due to the constraints in the second line of (5.107).

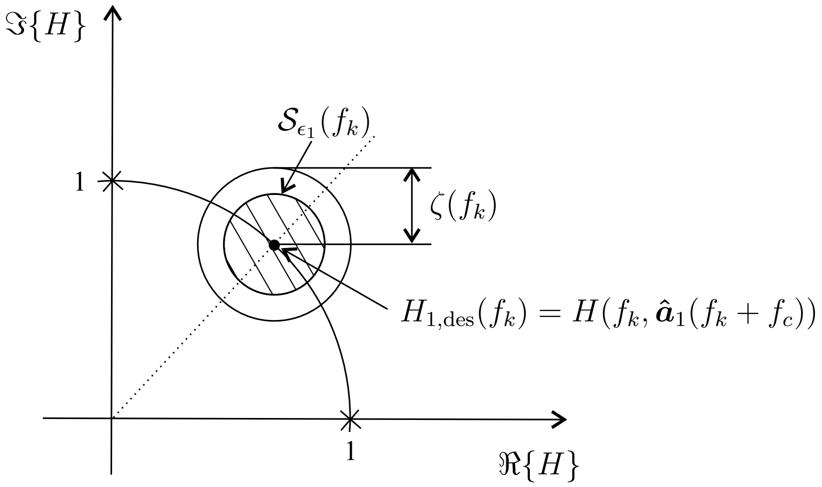


Figure 5.5. Geometric interpretation of the constraints of (5.107).

5.2.4 Relation to the norm-bounded broadband MVDR beamformer

In Section 5.2.3, we showed that adding the distortionless constraint (5.24) to the first and second proposed FIR beamformers leads to the same beamformer (5.107) if $\gamma(f_k)$ and $\zeta(f_k)$ are related as in (5.108). In this section, we study the relation of (5.107) to the norm-bounded broadband MVDR beamformer (5.37).

Towards this end, we write

$$\boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f_k) = \left(\boldsymbol{d}_{\mathrm{cs}}^T(f_k) \otimes \boldsymbol{I}_N\right) \underline{\boldsymbol{w}}^*, \tag{5.109}$$

where (5.28) has been used. Therefore, the constraints in the second line of (5.107) can be written as

$$\underline{\boldsymbol{w}}^H \left(\boldsymbol{\Pi}_{\boldsymbol{d}_{\mathrm{cs}}(f_k)} \otimes \boldsymbol{I}_N\right) \underline{\boldsymbol{w}} \leq \frac{\zeta^2(f_k)}{\epsilon_1^2(f_k + f_c) M_T}, \quad k = 1, \ldots, N_f, \tag{5.110}$$

where $\boldsymbol{\Pi}_{\boldsymbol{d}_{\mathrm{cs}}(f_k)}$ is the orthogonal projection matrix onto the 1D subspace spanned by $\boldsymbol{d}_{\mathrm{cs}}(f_k)$, and we used that [Bel60]

$$(\boldsymbol{M}_1 \otimes \boldsymbol{M}_2)(\boldsymbol{M}_3 \otimes \boldsymbol{M}_4) = (\boldsymbol{M}_1 \boldsymbol{M}_3) \otimes (\boldsymbol{M}_2 \boldsymbol{M}_4) \tag{5.111}$$

for any matrices $\boldsymbol{M}_1$, $\boldsymbol{M}_2$, $\boldsymbol{M}_3$, and $\boldsymbol{M}_4$ of conformable dimensions. Using (5.110), the beamformer (5.107) can be written as

$$\begin{aligned} \min_{\underline{\boldsymbol{w}}} \quad & \underline{\boldsymbol{w}}^H \underline{\hat{\boldsymbol{R}}_{\boldsymbol{x}}} \underline{\boldsymbol{w}} \\ \text{s.t.} \quad & \underline{\boldsymbol{w}}^H \left(\boldsymbol{\Pi}_{\boldsymbol{d}_{\mathrm{cs}}(f_k)} \otimes \boldsymbol{I}_N\right) \underline{\boldsymbol{w}} \leq \tfrac{\zeta^2(f_k)}{\epsilon_1^2(f_k+f_c) M_T}, \quad k = 1, \ldots, N_f \\ & \underline{\boldsymbol{C}}_{\mathrm{MVDR}}^H \underline{\boldsymbol{w}} = \boldsymbol{e}_{M_0}. \end{aligned} \tag{5.112}$$

The matrices

$$\mathbf{\Pi}_{\boldsymbol{d}_{\mathrm{cs}}(f_k)} \otimes \boldsymbol{I}_N, \quad k = 1, \ldots, N_f \tag{5.113}$$

are orthogonal projection matrices of rank N. Hence, the constraints in the second line of (5.112) limit the norm of the projection of $\underline{\boldsymbol{w}}$ onto different N-dimensional subspaces. In contrast, the beamformer (5.37) limits the Euclidean norm of $\underline{\boldsymbol{w}}$.

Depending on the choice of $\zeta(f_k)$ and the frequency samples f_k, the constraints in the second line of (5.112) may be more restrictive or less restrictive than the norm-constraint in (5.37). For example, (5.112) can be relaxed as

$$\begin{aligned} \min_{\underline{\boldsymbol{w}}} \quad & \underline{\boldsymbol{w}}^H \underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}} \underline{\boldsymbol{w}} \\ \text{s.t.} \quad & \underline{\boldsymbol{w}}^H \left(\sum_{k=1}^{N_f} \mathbf{\Pi}_{\boldsymbol{d}_{\mathrm{cs}}(f_k)} \otimes \boldsymbol{I}_N \right) \underline{\boldsymbol{w}} \le \tfrac{\zeta^2(f_k) N_f}{\epsilon_1^2(f_k + f_c) M_T} \\ & \underline{\boldsymbol{C}}_{\mathrm{MVDR}}^H \underline{\boldsymbol{w}} = \boldsymbol{e}_{M_0}. \end{aligned} \tag{5.114}$$

If the frequency samples are uniformly distributed over the frequency range as

$$f_k = -\frac{1}{2T_s} + \frac{k-1}{N_f T_s}, \quad k = 1, \ldots, N_f, \tag{5.115}$$

we have

$$\sum_{k=1}^{N_f} \mathbf{\Pi}_{\boldsymbol{d}_{\mathrm{cs}}(f_k)} = \frac{N_f}{M_T} \boldsymbol{I}_{M_T}. \tag{5.116}$$

In this case, (5.114) can be written as

$$\begin{aligned} \min_{\underline{\boldsymbol{w}}} \quad & \underline{\boldsymbol{w}}^H \underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}} \underline{\boldsymbol{w}} \\ \text{s.t.} \quad & \underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}} \le \tfrac{\zeta^2(f_k)}{\epsilon_1^2(f_k + f_c)} \\ & \underline{\boldsymbol{C}}_{\mathrm{MVDR}}^H \underline{\boldsymbol{w}} = \boldsymbol{e}_{M_0}. \end{aligned} \tag{5.117}$$

Hence, (5.37) is a relaxation of (5.107) if the frequency samples are chosen as in (5.115) and

$$\alpha_w \ge \frac{\zeta^2(f_k)}{\epsilon_1^2(f_k + f_c)}. \tag{5.118}$$

Similarly, it can be shown that (5.107) is a relaxation of (5.37) if the frequency samples are chosen as in (5.115) and

$$\alpha_w \le \frac{\zeta^2(f_k)}{\epsilon_1^2(f_k + f_c) N_f}. \tag{5.119}$$

5.2.5 Implementation without frequency discretization

The beamformers proposed in Sections 5.2.1 and 5.2.2 enforce constraints on the frequency response towards the desired signal only for a finite set of frequencies. To avoid the frequency discretization and associated errors, we modify in this section the beamforming problem (5.75), such that it can be solved without frequency discretization using the idea of [DLS02, GHNvD03, RV06, Dum07]. Similar steps allow to approximately solve all beamforming problems discussed in Sections 5.2.1 and 5.2.2 without frequency discretization.

Let us modify the beamformer of Section 5.2.1 as follows:

- First, let us use the 1-norm to limit the set of signal steering vectors instead of the 2-norm as in (5.70). To prevent the suppression of the desired signal, we use the constraints

  $$\begin{aligned} &\Re\left\{\mathbf{1}_N^T \boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f) + \tilde{\boldsymbol{\delta}}^T \boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f)\right\} \geq \breve{\gamma}(f) \\ &\forall\ \|\tilde{\boldsymbol{\delta}}\|_1 \leq \breve{\epsilon}_1(f+f_c),\ f \in [-B_s/2, B_s/2]\,, \end{aligned} \tag{5.120}$$

  where $\|\cdot\|_1$ is the vector 1-norm, and $\breve{\epsilon}_1(f+f_c)$ is an estimated upper bound on $\|\boldsymbol{\delta}_1(f+f_c)\|_1$. The constraints in (5.120) are satisfied if and only if

  $$\begin{aligned} \min_{\tilde{\boldsymbol{\delta}}} \quad & \Re\left\{\mathbf{1}_N^T \boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f) + \tilde{\boldsymbol{\delta}}^T \boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f)\right\} \\ \text{s.t.} \quad & \|\tilde{\boldsymbol{\delta}}\|_1 \leq \breve{\epsilon}_1(f+f_c) \end{aligned} \tag{5.121}$$

  is larger than or equal to $\breve{\gamma}(f)$ for all $f \in [-B_s/2, B_s/2]$. The worst-case mismatch vector of (5.121) can be expressed as

  $$\boldsymbol{\delta}_{1,\mathrm{wc}}(f+f_c) = -\frac{\boldsymbol{e}_{n_{\mathrm{wc}}}^T \boldsymbol{W} \boldsymbol{d}_{\mathrm{cs}}^*(f)}{|\boldsymbol{e}_{n_{\mathrm{wc}}}^T \boldsymbol{W} \boldsymbol{d}_{\mathrm{cs}}^*(f)|}\, \breve{\epsilon}_1(f+f_c)\boldsymbol{e}_{n_{\mathrm{wc}}}, \tag{5.122}$$

  where

  $$n_{\mathrm{wc}} = \arg \max_{n=1,\ldots,N} |\boldsymbol{e}_n^T \boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f)|. \tag{5.123}$$

  The solution of (5.121) follows as

  $$\Re\left\{\mathbf{1}_N^T \boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f)\right\} - \breve{\epsilon}_1(f+f_c)\|\boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f)\|_\infty, \tag{5.124}$$

  where $\|\cdot\|_\infty$ is the vector infinity-norm. This leads to the beamforming problem

  $$\begin{aligned} \min_{\boldsymbol{W}} \quad & \mathrm{vec}^H\{\boldsymbol{W}\}\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\mathrm{vec}\{\boldsymbol{W}\} \\ \text{s.t.} \quad & \Re\left\{\mathbf{1}_N^T \boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f)\right\} - \breve{\epsilon}_1(f+f_c)\|\boldsymbol{W}^* \boldsymbol{d}_{\mathrm{cs}}(f)\|_\infty \geq \breve{\gamma}(f) \\ & \forall\ f \in [-B_s/2, B_s/2]\,. \end{aligned} \tag{5.125}$$

- In the second step, let us replace the constraints in (5.125) by constraints that are linear in $\boldsymbol{W}$. The constraints in (5.125) can be written equivalently as

$$\begin{aligned}&\Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\text{cs}}(f)\right\}-\breve{\epsilon}_1(f+f_c)\left|\boldsymbol{e}_n^T\boldsymbol{W}^*\boldsymbol{d}_{\text{cs}}(f)\right|\geq\breve{\gamma}(f)\\&\forall\, n=1,\ldots,N,\quad f\in[-B_s/2,B_s/2]\,.\end{aligned}\tag{5.126}$$

To establish linearity in $\boldsymbol{W}$, we tighten the constraints in (5.126) as

$$\begin{aligned}&\Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\text{cs}}(f)\right\}-\breve{\epsilon}_1(f+f_c)\Re\left\{(k_1+jk_2)\boldsymbol{e}_n^T\boldsymbol{W}^*\boldsymbol{d}_{\text{cs}}(f)\right\}\geq\breve{\gamma}(f)\\&\forall\, k_1,k_2\in\{-1,1\},\quad n=1,\ldots,N,\quad f\in[-B_s/2,B_s/2]\,,\end{aligned}\tag{5.127}$$

where we exploited that for arbitrary complex numbers z

$$\begin{aligned}|z|&\leq|\,\Re\{z\}|+|\,\Im\{z\}|\\&=\max_{k_1,k_2\in\{-1,1\}}k_1\Re\{z\}+k_2\Im\{z\}\\&=\max_{k_1,k_2\in\{-1,1\}}\Re\{(k_1+jk_2)z\}.\end{aligned}\tag{5.128}$$

This leads to the beamforming problem

$$\begin{aligned}\min_{\boldsymbol{W}}\quad&\text{vec}^H\{\boldsymbol{W}\}\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\text{vec}\{\boldsymbol{W}\}\\\text{s.t.}\quad&\Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\text{cs}}(f)-\breve{\epsilon}_1(f+f_c)(k_1+jk_2)\boldsymbol{e}_n^T\boldsymbol{W}^*\boldsymbol{d}_{\text{cs}}(f)\right\}\geq\breve{\gamma}(f)\\&\forall\, k_1,k_2\in\{-1,1\},\quad n=1,\ldots,N,\quad f\in[-B_s/2,B_s/2]\,.\end{aligned}\tag{5.129}$$

The solution of (5.129) is in general not equal to the solution of (5.125). However, the optimum weight matrix of (5.129) satisfies the constraints of (5.125). If the optimum weight matrix of (5.125) satisfies

$$\Im\{\boldsymbol{e}_n^T\boldsymbol{W}^*\boldsymbol{d}_{\text{cs}}(f)\}=0\quad\forall\, n=1,\ldots,N,\quad f\in[-B_s/2,B_s/2]\,,\tag{5.130}$$

then (5.125) and (5.129) are equivalent. Note that (5.130) is equivalent to the presteering derivative constraints (5.36).

- In the third step, let us approximate $\breve{\epsilon}_1(f+f_c)$ and $\breve{\gamma}(f)$ as

$$\begin{aligned}\breve{\epsilon}_1(f+f_c)&\approx\boldsymbol{q}_{\breve{\epsilon}}^H\boldsymbol{d}_{\text{cs}}(f)\quad\forall\, f\in[-B_s/2,B_s/2]\\\breve{\gamma}(f)&\approx\boldsymbol{q}_{\breve{\gamma}}^H\boldsymbol{d}_{\text{cs}}(f)\quad\forall\, f\in[-B_s/2,B_s/2]\,,\end{aligned}\tag{5.131}$$

where $\boldsymbol{q}_{\breve{\epsilon}}$ and $\boldsymbol{q}_{\breve{\gamma}}$ are $M_T\times 1$ conjugate symmetric vectors, so that $\boldsymbol{q}_{\breve{\epsilon}}^H\boldsymbol{d}_{\text{cs}}(f)$ and $\boldsymbol{q}_{\breve{\gamma}}^H\boldsymbol{d}_{\text{cs}}(f)$ are real-valued. For notational simplicity, we expand $\breve{\epsilon}_1(f+f_c)$ and $\breve{\gamma}(f)$ using only trigonometric functions up to the order M_0-1. However, the proposed technique can be extended straightforwardly to higher-order trigonometric functions, which allows to improve the accuracy of (5.131).

The resulting beamforming problem can be expressed as

$$\begin{aligned} \min_{\boldsymbol{W}} \quad & \mathrm{vec}^H\{\boldsymbol{W}\}\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\mathrm{vec}\{\boldsymbol{W}\} \\ \text{s.t.} \quad & \Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f) - (k_1 + jk_2)\boldsymbol{d}_{\mathrm{cs}}^H(f)\boldsymbol{q}_{\breve{\epsilon}}\boldsymbol{e}_n^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f) - \boldsymbol{q}_{\breve{\gamma}}^H\boldsymbol{d}_{\mathrm{cs}}(f)\right\} \geq 0 \\ & \forall\, k_1, k_2 \in \{-1, 1\}, \quad n = 1, \ldots, N, \quad f \in [-B_s/2, B_s/2]\,. \end{aligned} \tag{5.132}$$

The problem (5.132) can be solved without frequency discretization. Towards this end, we use that

$$\mathcal{S}_{\boldsymbol{p}} = \left\{\boldsymbol{p} \in \mathbb{R} \times \mathbb{C}^{M_T-1} | \Re\left\{\boldsymbol{p}^H\boldsymbol{d}(f)\right\} \geq 0 \quad \forall\, f \in [-B_s/2, B_s/2]\right\} \tag{5.133}$$

is a convex set that can be represented by a finite number of linear matrix inequalities (LMIs) [DLS02]. Before specifying this LMI representation, let us first transform the inequality constraints of (5.132) into the form of (5.133).

For this purpose, we use that for arbitrary complex $M_T \times 1$ vectors $\boldsymbol{z}$, we have

$$\Re\left\{\boldsymbol{z}^H\boldsymbol{d}_{\mathrm{cs}}(f)\right\} = \Re\left\{\left(\boldsymbol{z}^H + \boldsymbol{z}^T\boldsymbol{J}_{M_T}\right)\underline{\boldsymbol{I}}\boldsymbol{d}(f)\right\}, \tag{5.134}$$

where

$$\underline{\boldsymbol{I}} = \begin{bmatrix} \mathbf{0}_{M_0-1\times M_T} \\ 1/2\,\boldsymbol{e}_1^T \\ \boldsymbol{e}_2^T \\ \vdots \\ \boldsymbol{e}_{M_0}^T \end{bmatrix}, \tag{5.135}$$

$\mathbf{0}_{m\times n}$ is the $m \times n$ matrix of zeroes, and $\boldsymbol{J}_{M_T}$ is the $M_T \times M_T$ exchange matrix, which contains ones on its anti-diagonal, and zeroes elsewhere. Furthermore, for arbitrary $M_T \times M_T$ complex matrices $\boldsymbol{M}$, we have

$$\Re\left\{\boldsymbol{d}_{\mathrm{cs}}^H(f)\boldsymbol{M}\boldsymbol{d}_{\mathrm{cs}}(f)\right\} = \frac{1}{2}\Re\left\{f_1^H(\boldsymbol{M} + \boldsymbol{M}^H)\boldsymbol{d}(f)\right\}, \tag{5.136}$$

where

$$\boldsymbol{f}_1(\boldsymbol{M}) = \begin{bmatrix} \mathrm{tr}\left(\boldsymbol{T}_0^{(M_T)}\boldsymbol{M}\right) \\ 2\,\mathrm{tr}\left(\boldsymbol{T}_1^{(M_T)}\boldsymbol{M}\right) \\ \vdots \\ 2\,\mathrm{tr}\left(\boldsymbol{T}_{M_T-1}^{(M_T)}\boldsymbol{M}\right) \end{bmatrix}, \tag{5.137}$$

and $\boldsymbol{T}_m^{(M_T)}$ is an $M_T \times M_T$ matrix defined as

$$\left[\boldsymbol{T}_m^{(M_T)}\right]_{k,l} = \begin{cases} 1 & \text{if } l = k+m, \\ 0 & \text{otherwise.} \end{cases} \tag{5.138}$$

Using (5.134) and (5.136), we obtain

$$\begin{aligned} & \Re\left\{\mathbf{1}_N^T \boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f) - (k_1 + jk_2)\boldsymbol{d}_{\text{cs}}^H(f)\boldsymbol{q}_{\breve{\epsilon}}\boldsymbol{e}_n^T \boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f) - \boldsymbol{q}_{\breve{\gamma}}^H \boldsymbol{d}_{\text{cs}}(f)\right\} \geq 0 \\ \Longleftrightarrow\ & \Re\big\{\big[\mathbf{1}_N^T \left(\boldsymbol{W}^* + \boldsymbol{W}\boldsymbol{J}_{M_T}\right)\underline{\boldsymbol{I}} - 2\boldsymbol{q}_{\breve{\gamma}}^H \underline{\boldsymbol{I}} \\ & - \boldsymbol{f}_1^H \left((k_1 + jk_2)\boldsymbol{q}_{\breve{\epsilon}}\boldsymbol{e}_n^T \boldsymbol{W}^* + (k_1 - jk_2)\boldsymbol{W}^T \boldsymbol{e}_n \boldsymbol{q}_{\breve{\epsilon}}^H\right)/2\big]\boldsymbol{d}(f)\big\} \geq 0. \end{aligned} \tag{5.139}$$

If $B_s = 1/T_s$, then the LMI representation of (5.133) can be expressed as [DLS02]

$$\mathcal{S}_{\boldsymbol{p}} = \left\{\boldsymbol{p} = \boldsymbol{f}_1(\boldsymbol{Y}) | \boldsymbol{Y} \in \mathcal{H}_+^{(M_T)}\right\}, \tag{5.140}$$

where $\mathcal{H}_+^{(M_T)}$ is the cone of positive semidefinite $M_T \times M_T$ Hermitian matrices. If $B_s < 1/T_s$, then the LMI representation of (5.133) is given by [DLS02]

$$\mathcal{S}_{\boldsymbol{p}} = \left\{\boldsymbol{p} = \boldsymbol{f}_1(\boldsymbol{Y}) - \boldsymbol{f}_2(\boldsymbol{Z}, B_s, T_s) | \boldsymbol{Y} \in \mathcal{H}_+^{(M_T)},\ \boldsymbol{Z} \in \mathcal{H}_+^{(M_T-1)}\right\}, \tag{5.141}$$

where

$$\boldsymbol{f}_2(\boldsymbol{Z}, B_s, T_s) = \begin{bmatrix} \Re\left\{\operatorname{tr}\left(\boldsymbol{Z}\left(\mu_1 \boldsymbol{T}_0^{(M_T-1)} + \mu_2 \boldsymbol{T}_1^{(M_T-1)}\right)\right)\right\} \\ \operatorname{tr}\left(\boldsymbol{Z}\left(2\mu_1 \boldsymbol{T}_1^{(M_T-1)} + \mu_2 \boldsymbol{T}_0^{(M_T-1)} + \mu_2 \boldsymbol{T}_2^{(M_T-1)}\right)\right) \\ \vdots \\ \operatorname{tr}\left(\boldsymbol{Z}\left(2\mu_1 \boldsymbol{T}_{M_T-3}^{(M_T-1)} + \mu_2 \boldsymbol{T}_{M_T-4}^{(M_T-1)} + \mu_2 \boldsymbol{T}_{M_T-2}^{(M_T-1)}\right)\right) \\ \operatorname{tr}\left(\boldsymbol{Z}\left(2\mu_1 \boldsymbol{T}_{M_T-2}^{(M_T-1)} + \mu_2 \boldsymbol{T}_{M_T-3}^{(M_T-1)}\right)\right) \\ \mu_2 \operatorname{tr}\left(\boldsymbol{Z}\boldsymbol{T}_{M_T-2}^{(M_T-1)}\right) \end{bmatrix}, \tag{5.142}$$

and

$$\begin{aligned} \mu_1 &= 2\cos(\pi B_s T_s)\left(1 - \cos(\pi B_s T_s)\right) \\ \mu_2 &= 2(\cos(\pi B_s T_s) - 1). \end{aligned} \tag{5.143}$$

For example, if $B_s < 1/T_s$, then (5.141) allows reformulating (5.132) as

$$\begin{aligned} \min_{\boldsymbol{W}, \{\boldsymbol{Y}_m\}, \{\boldsymbol{Z}_m\}} \quad & \operatorname{vec}^H\{\boldsymbol{W}\}\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}} \operatorname{vec}\{\boldsymbol{W}\} \\ \text{s.t.} \quad & \boldsymbol{f}_1(\boldsymbol{Y}_{(k_1+(k_2+3)/2)N+n}) - \boldsymbol{f}_2(\boldsymbol{Z}_{(k_1+(k_2+3)/2)N+n}, B_s, T_s) = \\ & \underline{\boldsymbol{I}}^T \left(\boldsymbol{W}^T + \boldsymbol{J}_{M_T}\boldsymbol{W}^H\right)\mathbf{1}_N - 2\underline{\boldsymbol{I}}^T \boldsymbol{q}_{\breve{\gamma}} \\ & \tfrac{1}{2}\boldsymbol{f}_1\left((k_1 + jk_2)\boldsymbol{q}_{\breve{\epsilon}}\boldsymbol{e}_n^T \boldsymbol{W}^* + (k_1 - jk_2)\boldsymbol{W}^T \boldsymbol{e}_n \boldsymbol{q}_{\breve{\epsilon}}^H\right), \\ & k_1, k_2 \in \{-1, 1\}, \quad n = 1, \ldots, N \\ & \boldsymbol{Y}_m \in \mathcal{H}_+^{(M_T)},\ \boldsymbol{Z}_m \in \mathcal{H}_+^{(M_T-1)}, \quad m = 1, \ldots, 4N. \end{aligned} \tag{5.144}$$

The optimization problem (5.144) is an SDP problem, which can be solved using, for example, the logarithmic barrier method [BV04]. In Appendix A.10, we show that the KKT system can be computed with the complexity $\mathcal{O}(M_T^2 N^2 + M_T^4 N)$. Moreover, we show that the KKT system can be solved with the complexity $\mathcal{O}(M_T^3 N^3 + M_T^6 N)$ by exploiting the block structure of the KKT matrix. Hence, for large FIR filter lengths, the computational complexity of (5.144) is higher than the complexity of (5.76).

5.3 DFT beamformers based on worst-case output power minimization

The DFT beamformers corresponding to the FIR beamformers proposed in Section 5.2 differ only with respect to the narrowband beamforming problems, which are solved for the subbands. We develop these narrowband beamforming problems in Sections 5.3.1 and 5.3.2. If $B_s < 1/T_s$, then some subbands lie outside the frequency range of the desired signal. The weight vectors of such subbands can be set to zero. In the following, we assume that $B_s = 1/T_s$, so that all subbands contain significant components of the desired signal.

5.3.1 Subband counterpart of the first proposed FIR beamformer

Similar steps as in Section 5.2.1 lead to the following narrowband beamforming problem for the qth subband:

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w} \\ \text{s.t.} \quad & \Re\{\boldsymbol{w}^H (\hat{\boldsymbol{a}}_1(f_q + f_c) + \tilde{\boldsymbol{\delta}})\} \geq \gamma(f_q) \\ & \forall \, \|\tilde{\boldsymbol{\delta}}\|_2 \leq \epsilon_1(f_q + f_c). \end{aligned} \tag{5.145}$$

The worst-case principle allows simplifying (5.145) as

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w} \\ \text{s.t.} \quad & \Re\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\} - \epsilon_1(f_q + f_c)\|\boldsymbol{w}\|_2 \geq \gamma(f_q). \end{aligned} \tag{5.146}$$

The latter optimization problem is of the same form like (4.31). Hence, (5.146) can be solved with the complexity $\mathcal{O}(N^3)$ using the technique of [LSW03]. The overall complexity of the DFT beamformer based on (5.146) is $\mathcal{O}(NQ \log Q + N^3 Q)$, where the first term results from the computation of the DFTs. This complexity is typically lower than the complexity of the corresponding FIR beamformer (5.76) if $M_T = Q$.

Similar as in Section 5.2.1.1, it can be shown that the constraint in (5.146) strictly limits the sensitivity to signal steering vector estimation errors. Furthermore, this constraint leads to an incentive for a low sensitivity if (5.81) holds. This incentive becomes stronger with increasing signal powers.

5.3.2 Subband counterpart of the second proposed FIR beamformer

The qth narrowband beamforming problem of the DFT beamformer corresponding to the second proposed FIR beamformer (5.90) is given by

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w} \\ \text{s.t.} \quad & |\boldsymbol{w}^H (\hat{\boldsymbol{a}}_1(f_q + f_c) + \tilde{\boldsymbol{\delta}}) - 1| \leq \zeta(f_q) \\ & \forall \, \|\tilde{\boldsymbol{\delta}}\|_2 \leq \epsilon_1(f_q + f_c). \end{aligned} \tag{5.147}$$

Similar steps as in Section 5.2.2 allow simplifying (5.147) as

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w} \\ \text{s.t.} \quad & |\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c) - 1| + \epsilon_1(f_q + f_c) \|\boldsymbol{w}\|_2 \leq \zeta(f_q), \end{aligned} \tag{5.148}$$

which is a SOCP problem. The latter optimization problem may appear as a novel narrowband beamforming approach. However, we show in the following that (5.148) is closely related to the norm-bounded MVDR beamformer (4.18) and the beamformer based on worst-case output power minimization (4.31).

Proposition 5.2. *The optimum weight vector of* (5.148) *satisfies*

$$|\boldsymbol{w}_q^H \hat{\boldsymbol{a}}_1(f_q + f_c) - 1| + \epsilon_1(f_q + f_c) \|\boldsymbol{w}_q\|_2 = \zeta(f_q), \tag{5.149}$$

$$\boldsymbol{w}_q^H \hat{\boldsymbol{a}}_1(f_q + f_c) = \Re \left\{ \boldsymbol{w}_q^H \hat{\boldsymbol{a}}_1(f_q + f_c) \right\}, \tag{5.150}$$

and

$$0 \leq \boldsymbol{w}_q^H \hat{\boldsymbol{a}}_1(f_q + f_c) \leq 1. \tag{5.151}$$

Proof. See Appendix A.11. □

Proposition 5.3. *The optimization problem* (5.148) *can be written equivalently as*

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w} \\ \text{s.t.} \quad & \Re\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\} - \epsilon_1(f_q + f_c)\|\boldsymbol{w}\|_2 \geq 1 - \zeta(f_q) \\ & \|\boldsymbol{w}\|_2 \leq \tfrac{\zeta(f_q)}{\epsilon_1(f_q+f_c)}. \end{aligned} \tag{5.152}$$

Proof. See Appendix A.12. □

To solve (5.152), we first solve the relaxed problem

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w} \\ \text{s.t.} \quad & \Re\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\} - \epsilon_1(f_q + f_c)\|\boldsymbol{w}\|_2 \geq 1 - \zeta(f_q). \end{aligned} \tag{5.153}$$

If the optimum weight vector of (5.153) is a feasible point of (5.152), then this weight vector also solves (5.152). Otherwise, the optimum weight vector of (5.152) has to satisfy the constraint in the last line of (5.152) with equality, i.e.,

$$\|\boldsymbol{w}_q\|_2 = \frac{\zeta(f_q)}{\epsilon_1(f_q + f_c)}. \tag{5.154}$$

Then, it follows from (5.149) that

$$\boldsymbol{w}_q^H \hat{\boldsymbol{a}}_1(f_q + f_c) = 1. \tag{5.155}$$

Adding the constraint $\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c) = 1$ to (5.152) yields

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w} \\ \text{s.t.} \quad & \boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c) = 1 \\ & \|\boldsymbol{w}\|_2 \leq \tfrac{\zeta(f_q)}{\epsilon_1(f_q+f_c)}, \end{aligned} \tag{5.156}$$

which has the same form as (4.18).

Note that the value of $\zeta(f_q)$ in (5.153) affects only the scaling of the optimum weight vector of this problem. Hence, the optimization problems (5.146) and (5.153) are equivalent in the sense that their optimum weight vectors are equal up to a scaling factor. Such a scaling factor is immaterial for the DFT beamformer output samples due to the scaling coefficients (5.53). Consequently, if the optimum weight vectors of (5.153) satisfy the constraints of (5.152) for all subbands, then the DFT beamformers based on (5.146) and (5.148) yield identical results.

Let $\tilde{\boldsymbol{w}}_q$ denote the optimum weight vector of (5.153). The constraint of (5.153) leads to an incentive for a small norm of $\tilde{\boldsymbol{w}}_q$ if (5.81) holds. If this incentive is strong enough such that the norm-constraint in (5.152) is satisfied for $\tilde{\boldsymbol{w}}_q$, then this vector solves (5.152). Otherwise, the solution of (5.152) follows from the norm-bounded MVDR beamforming problem (5.156).

The computational complexities of (5.153) and (5.156) are $\mathcal{O}(N^3)$ [Hud81, LSW03]. Hence, the order of the complexity of the second DFT beamformer is equal to that of the first DFT beamformer $\mathcal{O}(NQ\log Q + N^3 Q)$. Therefore, the complexities of the DFT beamformers are typically lower than the complexities of their FIR beamformer counterparts with $M_T = Q$.

5.4 Simulation results

We consider $L = 3$ far-field sources that impinge on a uniform linear array of $N = 5$ frequency-flat omnidirectional sensors. The sensors are aligned along the z-axis of the coordinate system. The distance between adjacent sensors is $c/(2f_c + B_s)$, where the carrier frequency is $f_c = 2$, and the signal bandwidth is $B_s = 1$.

The baseband signal and noise waveforms are modeled as standard complex Gaussian random processes. The baseband power spectral densities of the first and second sources are constant over the frequency interval $[-0.5, 0.5]$. The power spectral density of the third source is constant over the interval $[0, 0.25]$, and zero outside this interval. The INRs of both interferers are set to 30 dB.

The nth component of the presumed steering vector of the desired signal is

$$[\hat{\boldsymbol{a}}_1(f + f_c)]_n = e^{j2\pi(f+f_c)p_{z_n}\sin(\hat{\theta}_1)/c}, \tag{5.157}$$

where $\hat{\theta}_1 = -20°(\pi/180°)$. In each Monte-Carlo run, we randomly generate the components of the actual (true) signal steering vector as

$$[\boldsymbol{a}_1(f + f_c)]_n = (1 + g_n)e^{j2\pi(f+f_c)p_{z_n}\sin(\hat{\theta}_1+\Delta\theta)/c}, \quad n = 1, \dots, N, \tag{5.158}$$

where g_n and $\Delta\theta$ are zero-mean Gaussian random variables with the standard deviations 0.1 and $1.5°(\pi/180°)$, respectively. Using the same values of g_n, the steering vector components of the two interferers are

$$[\boldsymbol{a}_l(f + f_c)]_n = (1 + g_n)e^{j2\pi(f+f_c)p_{z_n}\sin(\theta_l)/c}, \quad l = 2, 3, \quad n = 1, \dots, N, \tag{5.159}$$

where $\theta_2 = 30°(\pi/180°)$, and $\theta_3 = 10°(\pi/180°)$. Throughout our simulations, $\epsilon_1(f_k+f_c)$ and $\breve{\epsilon}_1(f+f_c)$ are chosen so that the events

$$\|\boldsymbol{\delta}_1(f+f_c)\|_2 \le \epsilon_1(f+f_c) \quad \text{and} \quad \|\boldsymbol{\delta}_1(f+f_c)\|_1 \le \breve{\epsilon}_1(f+f_c) \tag{5.160}$$

occur with probability 0.99.

To simplify the comparison of FIR and DFT beamformers, we use presteering delays also for the DFT beamformers. We set $\bar{T}_p = 0$, the length of the FIR filters is $M_T = 5$, and the sampling period is $T_s = 1$. All beamformers based on a frequency discretization use the uniform frequency grid (5.49). Unless specified otherwise, we set $N_f = Q = 15$. The default value for the number of training snapshot vectors $\boldsymbol{x}(kT_s)$ is $K = 250$.

We compare the performance of the following beamformers:

- The first proposed FIR beamformer (5.76) with

$$\gamma(f_k) = 1, \quad k = 1, \ldots, N_f. \tag{5.161}$$

- The second proposed FIR beamformer (5.90) with

$$\zeta(f_k) = \xi \frac{\epsilon_1(f_k+f_c)}{\sqrt{N}}, \quad k = 1, \ldots, N_f. \tag{5.162}$$

  If not specified otherwise, we choose ξ such that $20\log_{10}\xi = 3$ dB throughout our simulations.

- The proposed FIR beamformer without frequency discretization (5.144) using

$$\breve{\gamma}(f) = 1 \quad \forall f \in \left[-\frac{B_s}{2}, \frac{B_s}{2}\right]. \tag{5.163}$$

- The beamformer of [EKKG05b].
- The DFT beamformer counterpart of the first proposed FIR beamformer with $\gamma(f_q)$ chosen as in (5.161).
- The DFT beamformer counterpart of the second proposed FIR beamformer with $\zeta(f_q)$ chosen as in (5.162).
- The norm-bounded MVDR beamformer (5.37) with $\alpha_w = \xi^2/N$.
- The norm-bounded MVDR beamformer (5.37) with $\alpha_w = \xi^2/N$ and the additional presteering derivative constraints (5.36).

To solve the proposed FIR beamforming problems, we use the MATLAB package `CVX` [GB08a, GB08b].

We show results for the following three performance measures. The distortion of the desired signal is measured in terms of the normalized signal distortion (NSD). The suppression of interferers and noise is measured in terms of the OSINR. The overall beamformer performance is measured in terms of the normalized mean squared error (NMSE). A good overall beamformer performance requires that the signal distortions are small and that interferers and noise are suppressed effectively. In the following, we define the NSD, OSINR, and NMSE for the FIR beamformers. Based on (5.65), the performance measures can be defined similarly for the DFT beamformers.

For the FIR beamformers, the NSD is defined as

$$\mathrm{NSD} \triangleq \frac{1}{P_1}\mathrm{E}\left\{\left|\underline{\boldsymbol{w}}^H\underline{\boldsymbol{x}}_1(kT_s) - s_1((k-M_0+1)T_s)\right|^2\right\}, \tag{5.164}$$

where $s_1(kT_s)$ is the complex envelope of the signal-of-interest at the kth time instant, $P_1 = \mathrm{E}\left\{|s_1(kT_s)|^2\right\}$ is the power of the signal-of-interest, and the beamformer weight vectors are scaled so that the NSD is minimized. The OSINR is defined as [vT02]

$$\mathrm{OSINR} \triangleq \frac{\underline{\boldsymbol{w}}^H\underline{\boldsymbol{R}}_1\underline{\boldsymbol{w}}}{\underline{\boldsymbol{w}}^H\underline{\boldsymbol{R}}_{i+n}\underline{\boldsymbol{w}}}, \tag{5.165}$$

where

$$\underline{\boldsymbol{R}}_{i+n} = \sum_{l=2}^{L}\underline{\boldsymbol{R}}_l + \underline{\boldsymbol{R}}_{\boldsymbol{n}} \tag{5.166}$$

is the $M_TN \times M_TN$ spatio-temporal interference-plus-noise covariance matrix. The optimum OSINR is given by [vT02]

$$\mathrm{OSINR}_{\mathrm{opt}} = \mathcal{P}(\underline{\boldsymbol{R}}_{i+n}^{-1}\underline{\boldsymbol{R}}_1), \tag{5.167}$$

where $\mathcal{P}(\cdot)$ denotes the principal eigenvalue operator. The overall beamformer performance is measured in terms of the NMSE, defined as

$$\mathrm{NMSE} \triangleq \frac{1}{P_1}\mathrm{E}\left\{\left|\underline{\boldsymbol{w}}^H\underline{\boldsymbol{x}}(kT_s) - s_1((k-M_0+1)T_s)\right|^2\right\}, \tag{5.168}$$

where we use the same scaling of the beamformer weight vectors as for the computation of the NSD. The NMSE is lower bounded by

$$\mathrm{NMSE}_{\mathrm{min}} = 1 - \frac{\underline{\boldsymbol{r}}_1^H\underline{\boldsymbol{R}}_{\boldsymbol{x}}^{-1}\underline{\boldsymbol{r}}_1}{P_1}, \tag{5.169}$$

where

$$\underline{\boldsymbol{r}}_1 = \mathrm{E}\left\{\underline{\boldsymbol{x}}(kT_s)s_1^*((k-M_0+1)T_s)\right\} = \int\limits_{-1/(2T_s)}^{1/(2T_s)} S_1(f)\underline{\boldsymbol{a}}_1(f)e^{j2\pi f(M_0-1)T_s}\mathrm{d}f. \qquad (5.170)$$

Figures 5.6–5.8 show the beamformer performance versus the input SNR. Figure 5.6 compares the performance of the FIR beamformers based on worst-case output power minimization, i.e., the first and second proposed FIR beamformers, the proposed FIR beamformer without frequency discretization, and the beamformer of [EKKG05b]. Figure 5.7 compares the performance of the first and second proposed FIR beamformers with the performance of their DFT beamformer counterparts. Figure 5.8 shows the performance of the first proposed FIR beamformer and the norm-bounded broadband MVDR beamformers with and without presteering derivative constraints.

Figures 5.9–5.11 depict the beamformer performance versus the number of training snapshot vectors $\boldsymbol{x}(kT_s)$. The SNR is set to 20 dB. All other parameters are chosen as before.

In Figure 5.12–5.14, we vary the presumed DOA of the desired signal and show the performance of the beamformers versus the angular distance $\theta_3 - \hat{\theta}_1$.

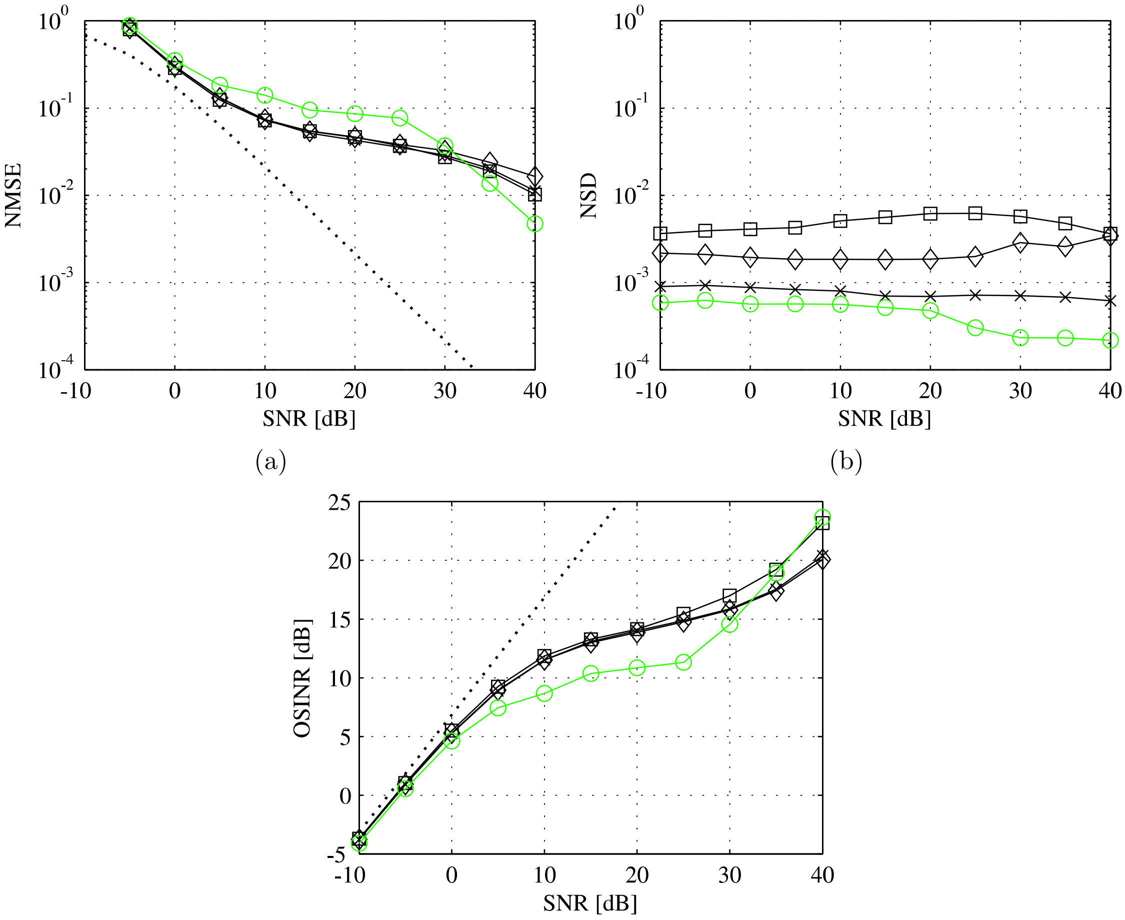
Figure 5.15 depicts the performance of the second proposed FIR beamformer, its DFT beamformer counterpart, and the norm-bounded broadband MVDR beamformers with and without presteering derivative constraints versus $20\log_{10}\xi$.

Figure 5.16 shows the performance of the beamformers based on a frequency discretization versus the number of frequency steps $N_f = Q$. The frequency samples are uniformly distributed over the frequency range as in (5.49).

The Figures 5.6–5.16 demonstrate that the proposed beamformers are attractive alternatives to the state-of-the-art robust broadband beamformers.

The NMSE performance of the beamformers is typically substantially worse than the NSD performance. Hence, the overall beamformer performance is usually dominated by the capability to suppress interferers and noise.

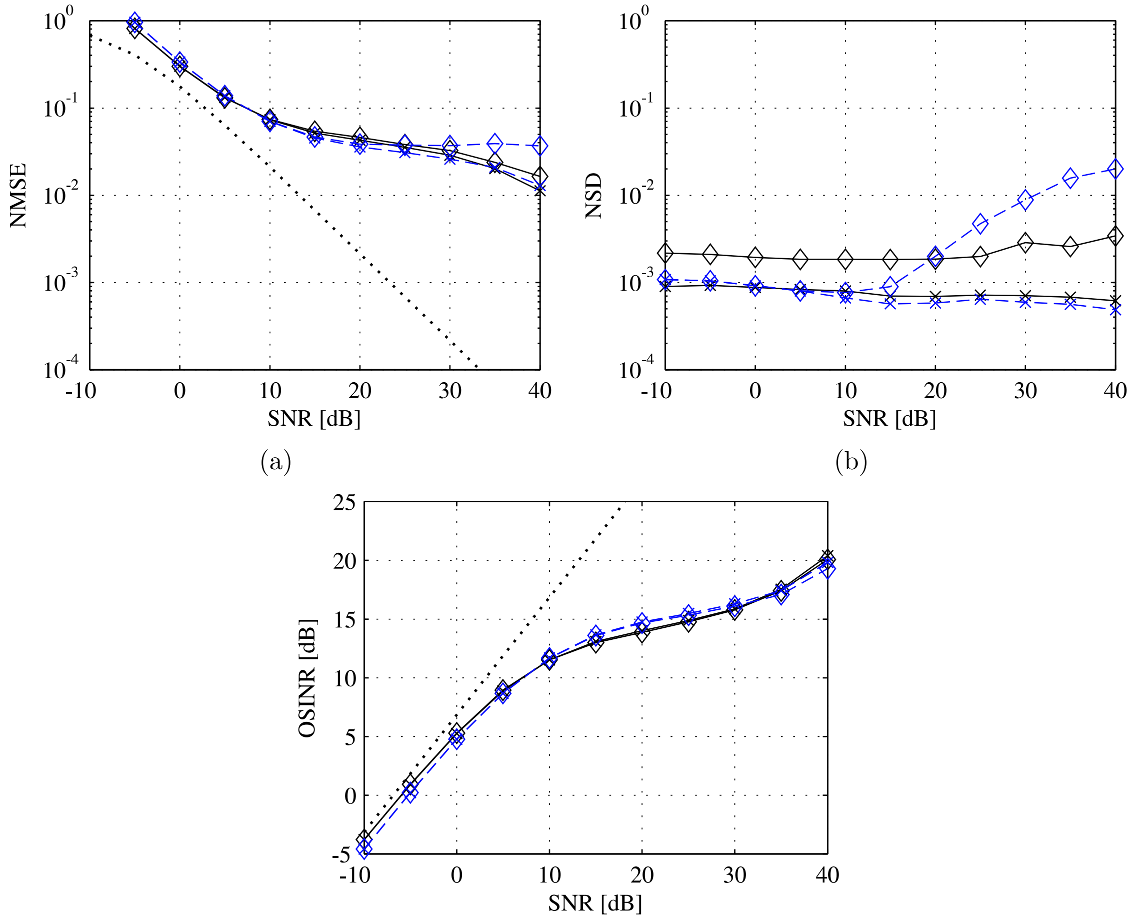
The beamformer of [EKKG05b] uses presteering derivative constraints to prevent phase distortions of the desired signal. These constraints are rather restrictive, as they decrease the beamformer degrees of freedom by a factor of two. The beamformer



(a) (b)

(c)

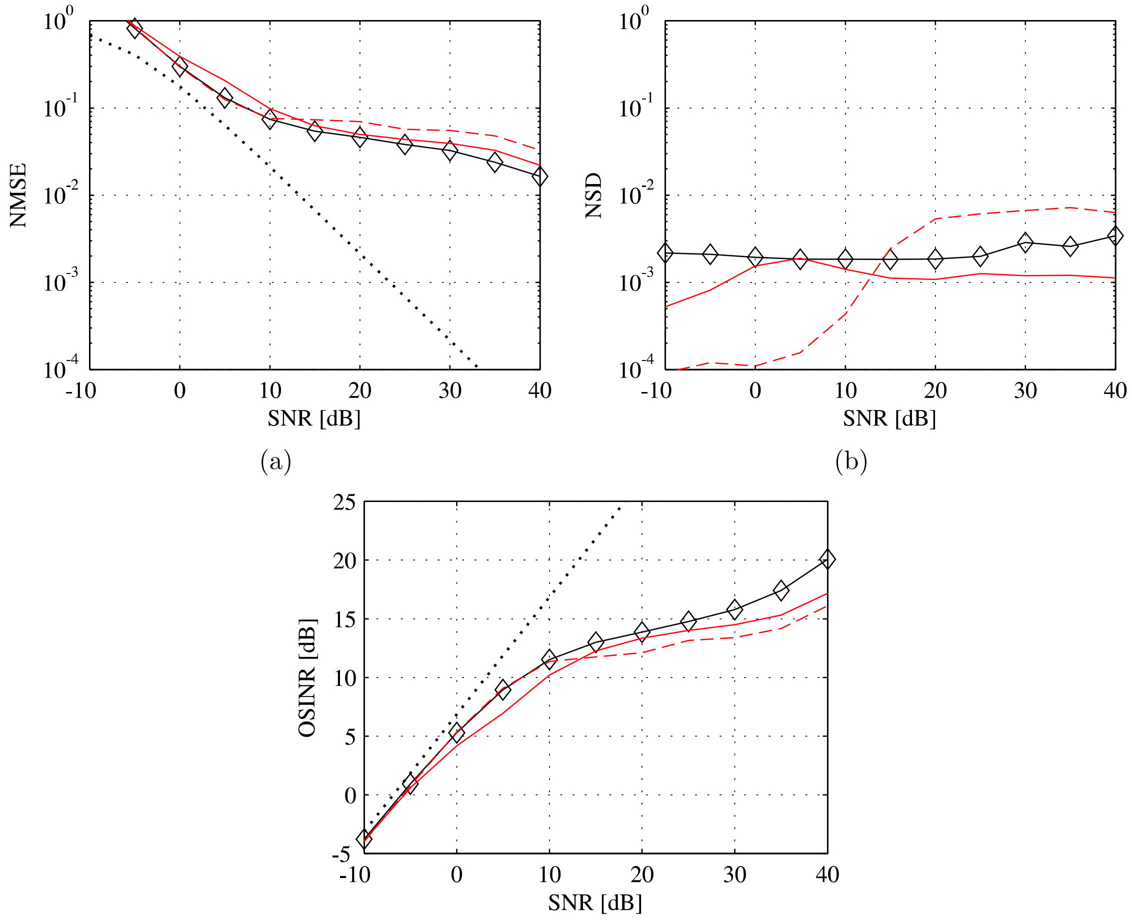
Figure 5.6. Performance versus input SNR. Solid line with diamonds: First proposed FIR beamformer. Solid line with crosses: Second proposed FIR beamformer. Solid line with squares: Proposed FIR beamformer without frequency discretization. Solid line with circles: Beamformer of [EKKG05b]. Dotted line: Minimum NMSE or optimum OSINR.



(a) (b)

(c)

Figure 5.7. Performance versus input SNR. Solid line with diamonds: First proposed FIR beamformer. Solid line with crosses: Second proposed FIR beamformer. Dashed line with diamonds: DFT beamformer counterpart of the first proposed FIR beamformer. Dashed line with crosses: DFT beamformer counterpart of the second proposed FIR beamformer. Dotted line: Minimum NMSE or optimum OSINR for the FIR beamformers.

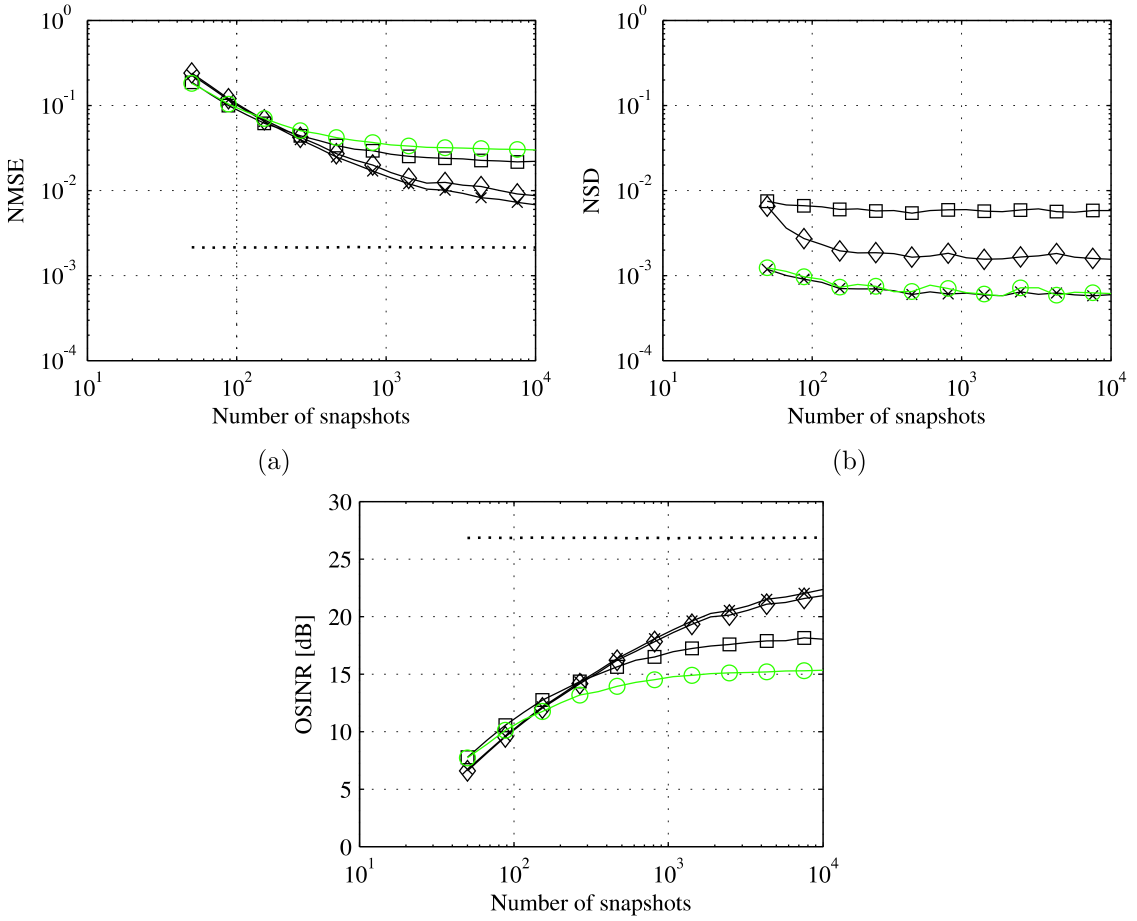


(a)

(b)

(c)

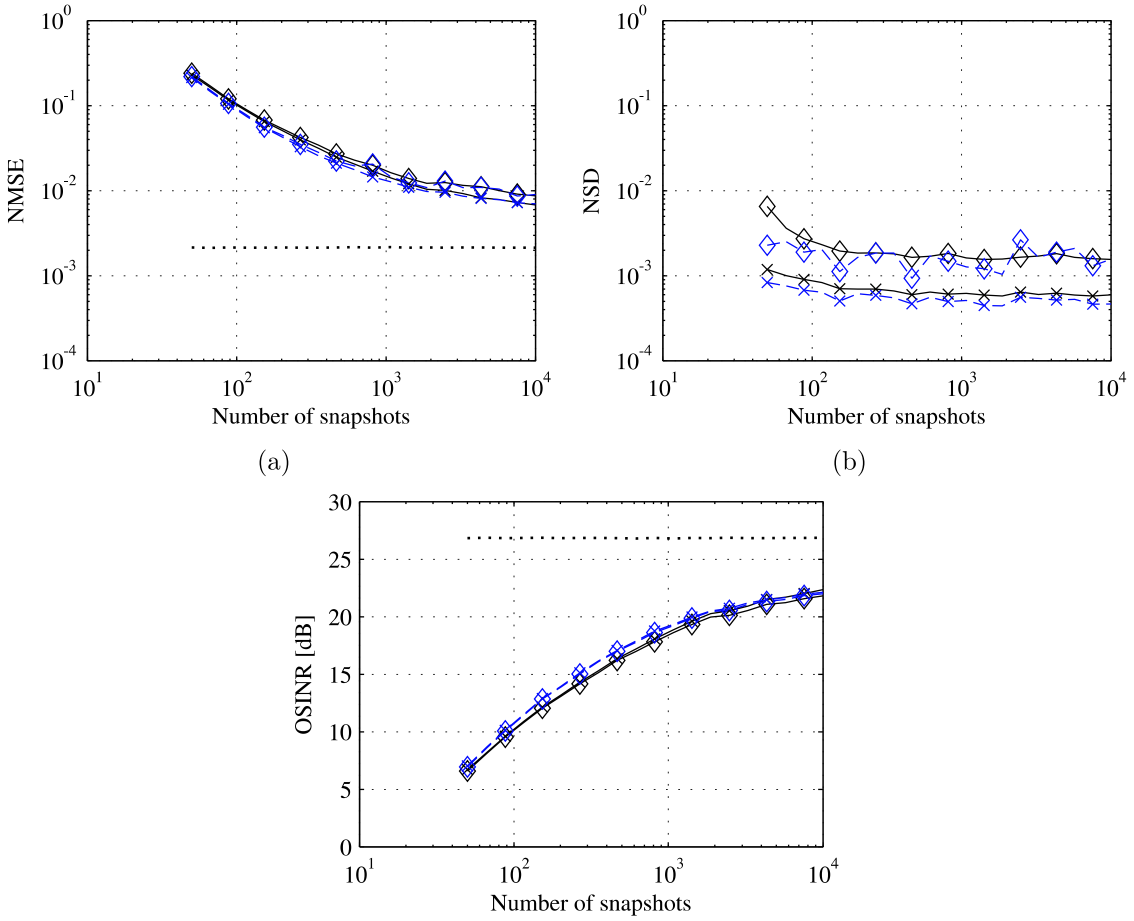
Figure 5.8. Performance versus input SNR. Solid line with diamonds: First proposed FIR beamformer. Solid line: Norm-bounded broadband MVDR beamformer. Dashed line: Norm-bounded broadband MVDR beamformer with presteering derivative constraints. Dotted line: Minimum NMSE or optimum OSINR.



(a) (b)

(c)

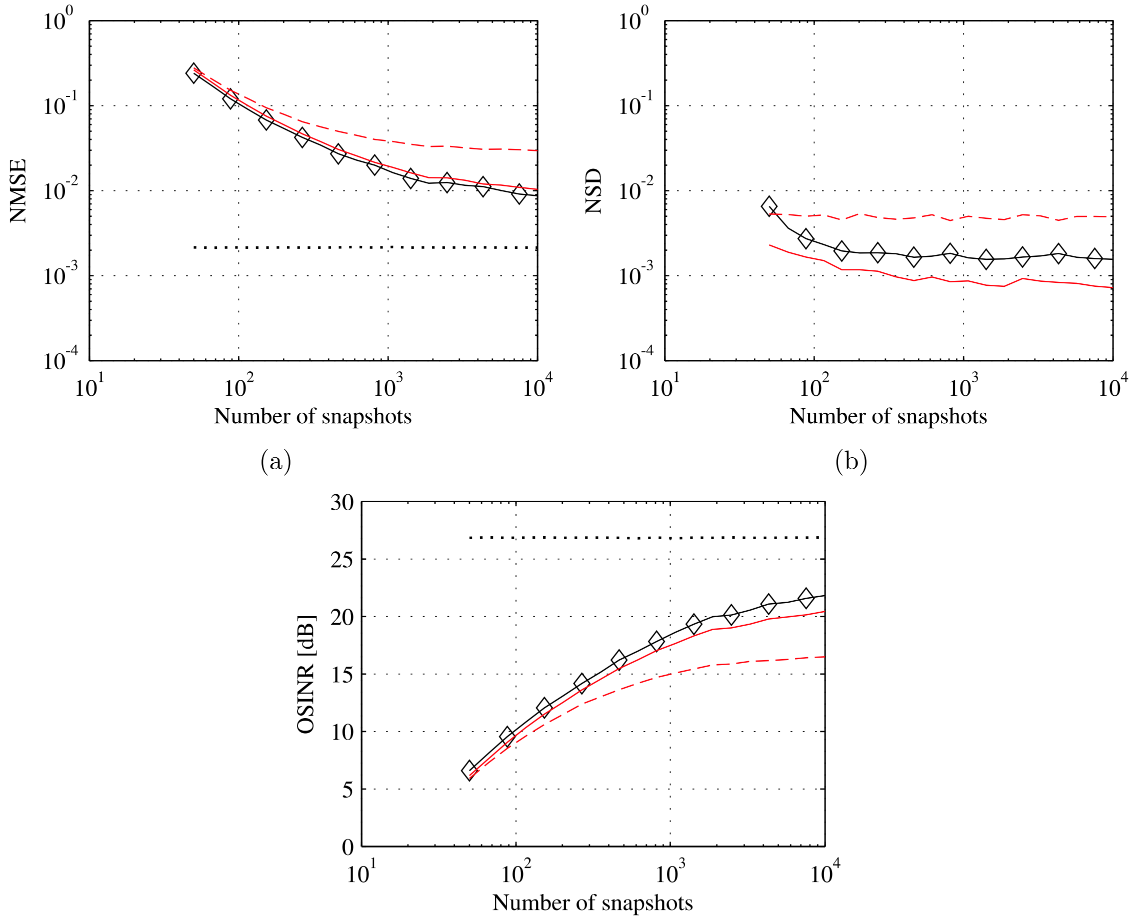
Figure 5.9. Performance versus number of snapshots. Solid line with diamonds: First proposed FIR beamformer. Solid line with crosses: Second proposed FIR beamformer. Solid line with squares: Proposed FIR beamformer without frequency discretization. Solid line with circles: Beamformer of [EKKG05b]. Dotted line: Minimum NMSE or optimum OSINR.



(a) (b)

(c)

Figure 5.10. Performance versus number of snapshots. Solid line with diamonds: First proposed FIR beamformer. Solid line with crosses: Second proposed FIR beamformer. Dashed line with diamonds: DFT beamformer counterpart of the first proposed FIR beamformer. Dashed line with crosses: DFT beamformer counterpart of the second proposed FIR beamformer. Dotted line: Minimum NMSE or optimum OSINR for the FIR beamformers.

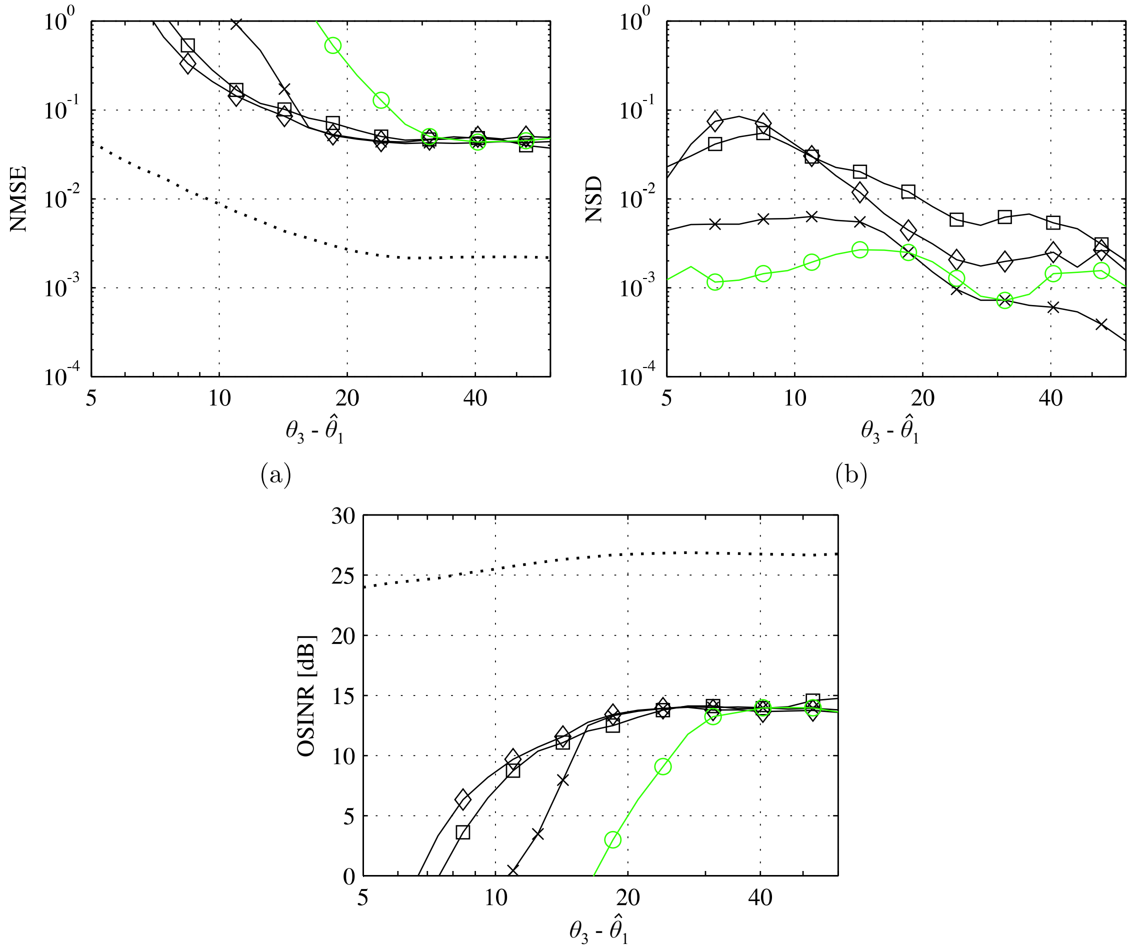


(a)

(b)

(c)

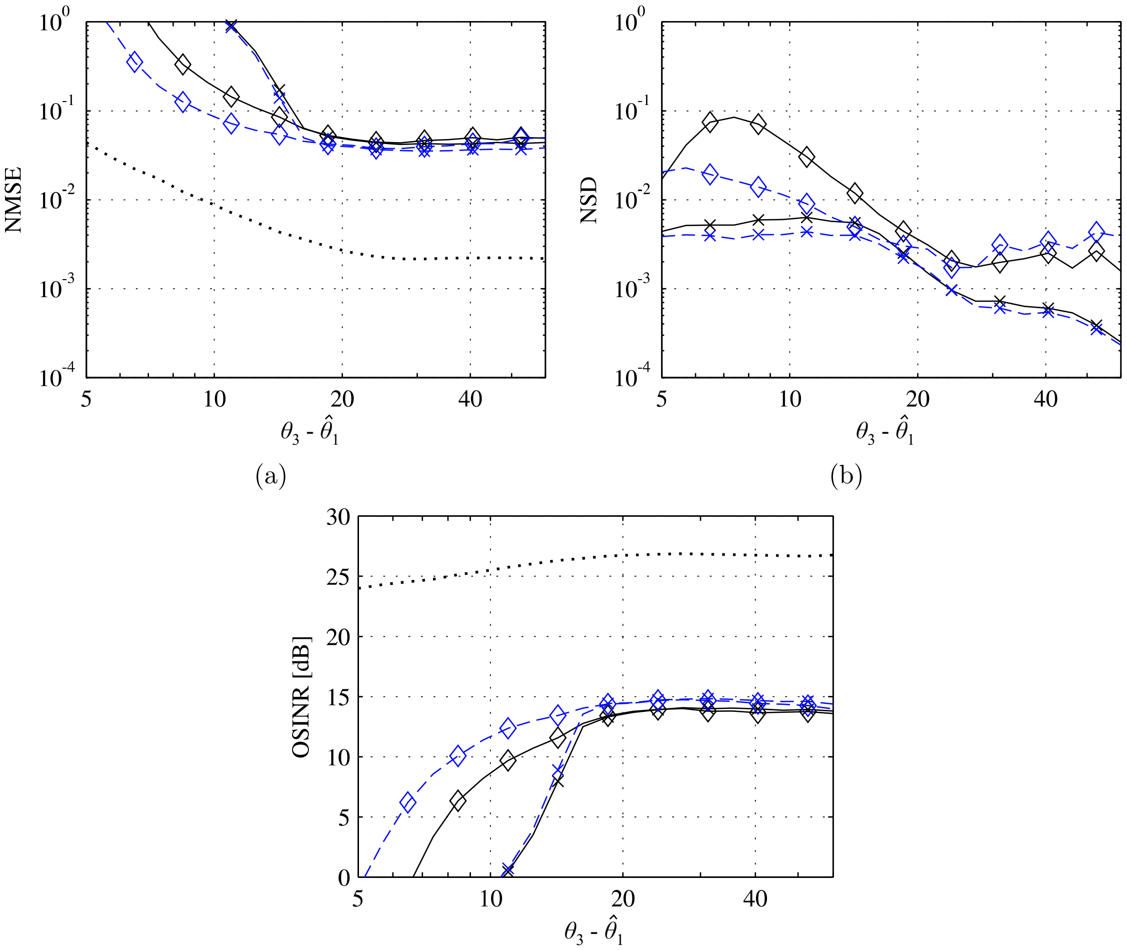
Figure 5.11. Performance versus number of snapshots. Solid line with diamonds: First proposed FIR beamformer. Solid line: Norm-bounded broadband MVDR beamformer. Dashed line: Norm-bounded broadband MVDR beamformer with presteering derivative constraints. Dotted line: Minimum NMSE or optimum OSINR.



(a) (b)

(c)

Figure 5.12. Performance versus $\theta_3 - \hat{\theta}_1$. Solid line with diamonds: First proposed FIR beamformer. Solid line with crosses: Second proposed FIR beamformer. Solid line with squares: Proposed FIR beamformer without frequency discretization. Solid line with circles: Beamformer of [EKKG05b]. Dotted line: Minimum NMSE or optimum OSINR.

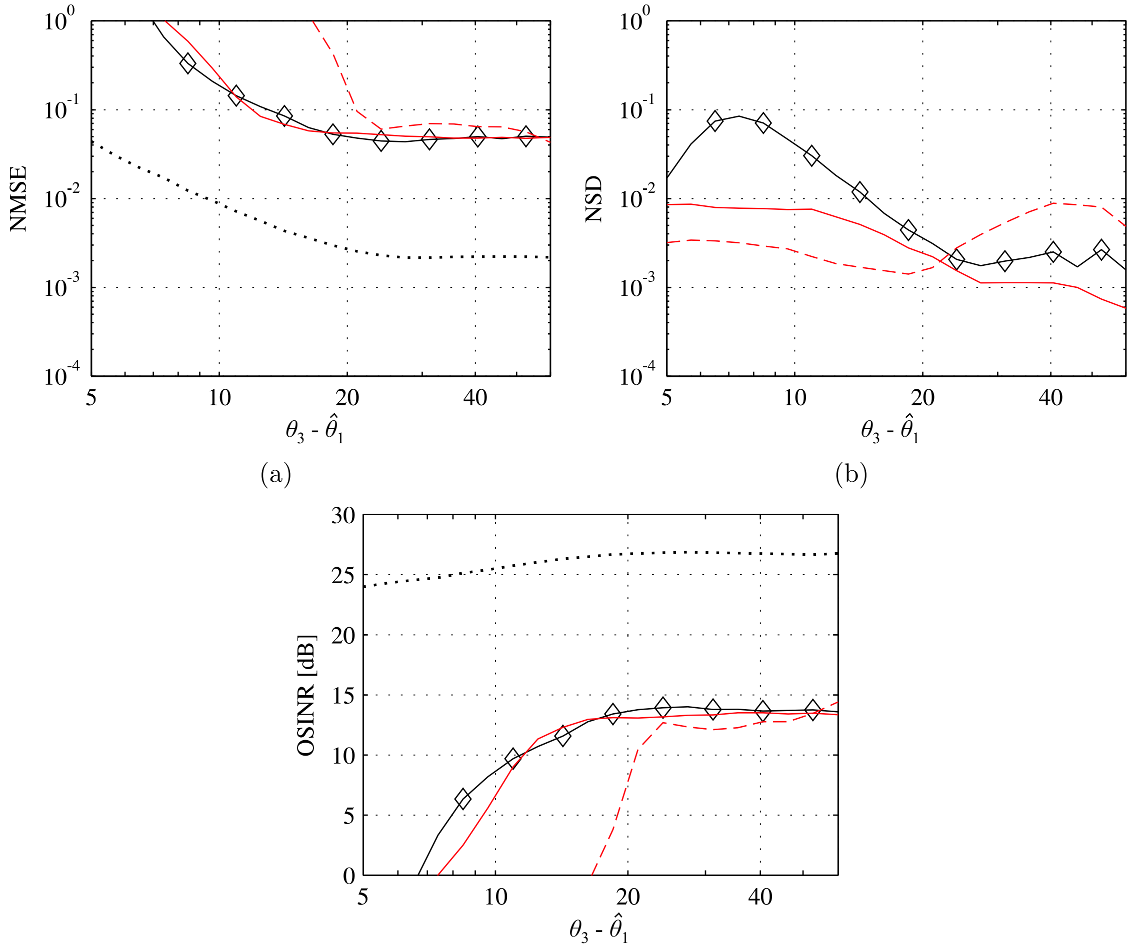


(a)

(b)

(c)

Figure 5.13. Performance versus $\theta_3 - \hat{\theta}_1$. Solid line with diamonds: First proposed FIR beamformer. Solid line with crosses: Second proposed FIR beamformer. Dashed line with diamonds: DFT beamformer counterpart of the first proposed FIR beamformer. Dashed line with crosses: DFT beamformer counterpart of the second proposed FIR beamformer. Dotted line: Minimum NMSE or optimum OSINR for the FIR beamformers.

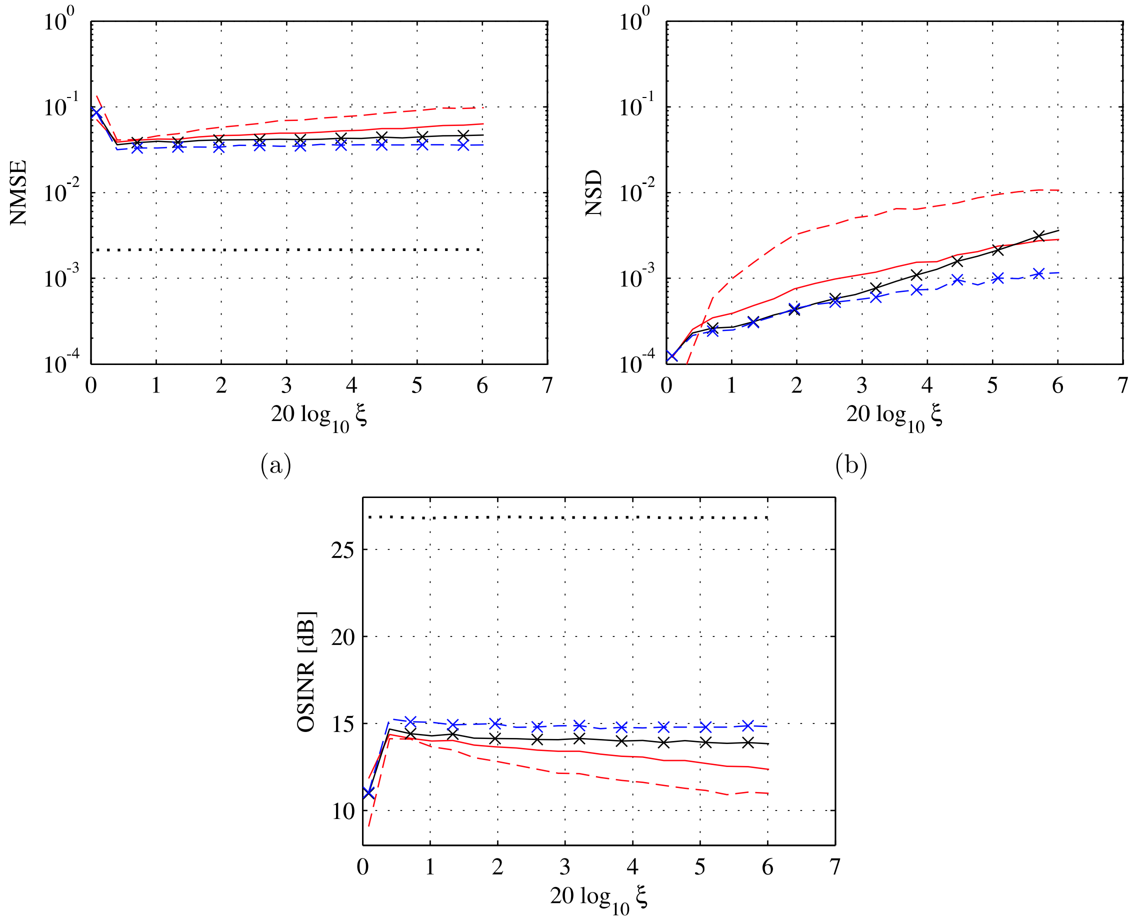


(a)

(b)

(c)

Figure 5.14. Performance versus $\theta_3 - \hat{\theta}_1$. Solid line with diamonds: First proposed FIR beamformer. Solid line: Norm-bounded broadband MVDR beamformer. Dashed line: Norm-bounded broadband MVDR beamformer with presteering derivative constraints. Dotted line: Minimum NMSE or optimum OSINR.

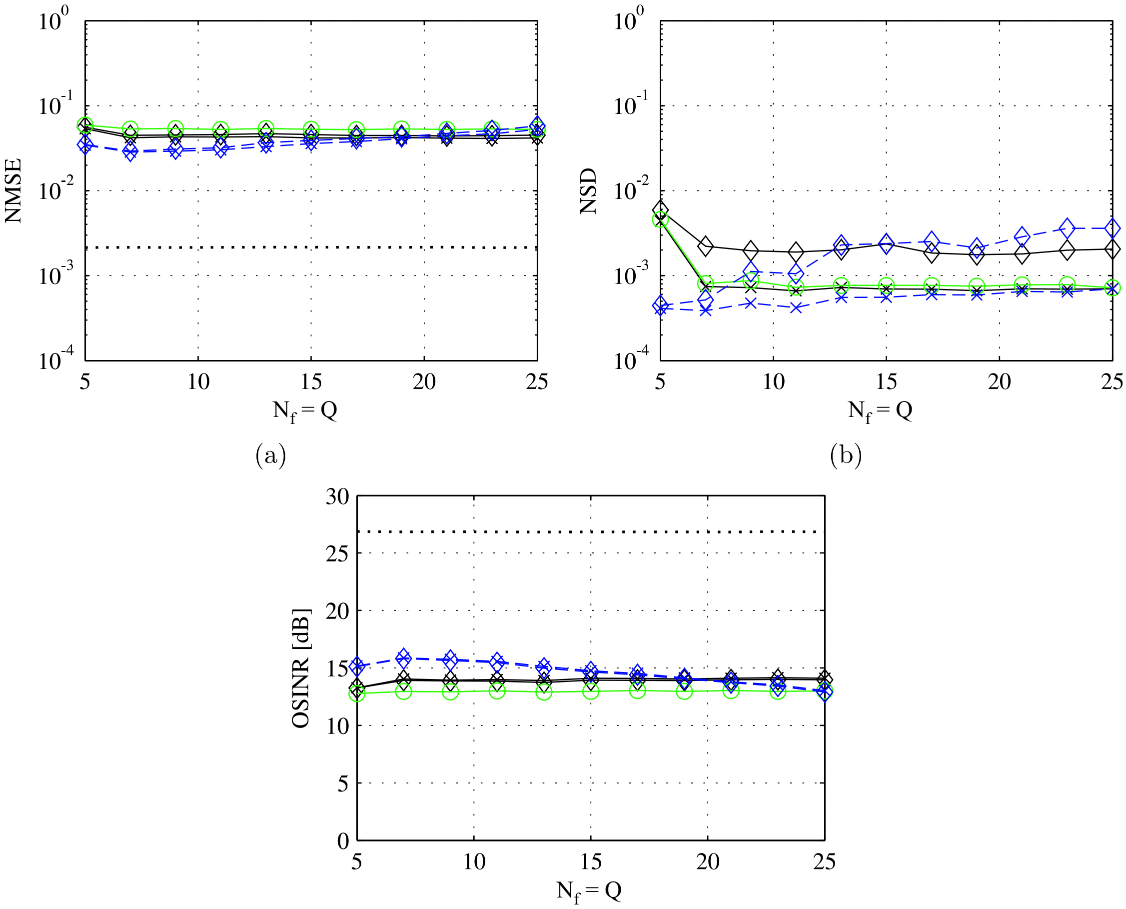


(a)

(b)

(c)

Figure 5.15. Performance versus ξ. Solid line with crosses: Second proposed FIR beamformer. Dashed line with crosses: DFT beamformer counterpart of the second proposed FIR beamformer. Solid line: Norm-bounded broadband MVDR beamformer. Dashed line: Norm-bounded broadband MVDR beamformer with presteering derivative constraints. Dotted line: Minimum NMSE or optimum OSINR for the FIR beamformers.



(a) (b)

(c)

Figure 5.16. Performance versus $N_f = Q$. Solid line with diamonds: First proposed FIR beamformer. Solid line with crosses: Second proposed FIR beamformer. Solid line with circles: Beamformer of [EKKG05b]. Dashed line with diamonds: DFT beamformer counterpart of the first proposed FIR beamformer. Dashed line with crosses: DFT beamformer counterpart of the second proposed FIR beamformer. Dotted line: Minimum NMSE or optimum OSINR for the FIR beamformers.

of [EKKG05b] and the norm-bounded broadband MVDR beamformer with presteering derivative constraints therefore have a rather poor capability to suppress interferers and noise, in particular in adverse environments with closely spaced sources, see Figures 5.12 and 5.14.

Figure 5.15 shows that our default value for ξ is non-optimum. If $20 \log_{10} \xi$ decreases below 0.4 dB, then the NMSE performance of all beamformers decreases substantially due to the decreasing capability to suppress interferers and noise. It can be expected that this threshold depends on the angular distance between the source-of-interest and its closest interferer. In particular, if an interferer is closely spaced to the source-of-interest, then a large value of ξ is required to effectively suppress this interferer. Figure 5.15 also verifies that distortions of the desired signal become stronger if ξ increases.

Figure 5.16 demonstrates that the number of frequency steps $N_f = Q$ can be chosen rather small without a significant degradation of the beamformer performance. This is beneficial for the computationally efficient implementation of the proposed FIR beamformers based on a frequency discretization.

5.5 Summary

In this chapter, we proposed two FIR beamformers based on worst-case output power minimization, which use different constraints to control the frequency response towards the desired signal. Similar to the norm-bounded broadband MVDR beamformer, the proposed FIR beamformers strictly limit the sensitivity to signal steering vector estimation errors. Additionally, the constraints of the proposed FIR beamformers lead to an incentive for a low sensitivity if the signal steering vectors lie within the presumed uncertainty sets. This incentive becomes stronger with increasing signal powers. Such an incentive does not exist for the norm-bounded broadband MVDR beamformer.

The proposed FIR beamforming problems are SOCP problems, which can be solved, for example, by means of the logarithmic barrier method. We showed that, for both proposed FIR beamformers, the barrier function, its gradient, and its Hessian can be computed with the complexity $\mathcal{O}(M_T^2 N^2 + M_T N^2 N_f)$, where M_T is the FIR filter length, N the number of sensors, and N_f the number of discrete frequencies, at which the beamformer response is controlled. The complexity for the exact computation of a Newton step is $\mathcal{O}((M_T N + N_f)^3)$ due to the inversion of the Hessian matrix. Therefore, the computational complexities of the proposed FIR beamformers are comparable to the complexity of the norm-bounded broadband MVDR beamformer.

We also formulated and analyzed the computationally efficient DFT beamformer counterparts of the proposed FIR beamformers. These DFT beamformers are based on the narrowband beamformer of [VGL03]. The two DFT beamformers differ in that the narrowband beamforming problems of one DFT beamformer contain an additional norm-constraint.

The proposed FIR beamformers control the frequency response towards the desired signal only for a finite set of frequencies. To avoid the frequency discretization and associated errors, we modified the first proposed FIR beamforming problem, such that it can be solved without frequency discretization. These modifications involve several approximations and result in an SDP problem, which can be solved by means of the logarithmic barrier method. We showed that the complexity to compute the corresponding KKT system is $\mathcal{O}(M_T^2 N^2 + M_T^4 N)$, and that this system can be solved with the complexity $\mathcal{O}(M_T^3 N^3 + M_T^6 N)$. Hence, for large FIR filter lengths, the computational complexity of this beamforming problem is higher than the complexities of the proposed FIR beamforming problems based on a frequency discretization.

Our simulation results demonstrate that the proposed beamformers are attractive alternatives to the state-of-the-art broadband beamformers. In particular, the proposed FIR beamformers offer a substantially improved performance as compared to previous FIR beamformers based on worst-case output power minimization.

Chapter 6

Conclusions and outlook

We have developed novel concepts and methods for DOA estimation, array geometry design, and beamforming.

- Two high-resolution rooting-based DOA estimation methods have been proposed, which can be applied to arbitrary array geometries. The proposed estimators provide attractive tradeoffs between DOA estimation performance and computational complexity.
- A novel array geometry design has been proposed for azimuthal DOA estimation. This array design is closely related to the MRA concept, but the sensors are not required to lie on a uniform grid. The proposed arrays provide 360° angular coverage. Simulation results suggest that the apertures of the proposed arrays increase quadratically with the number of sensors.
- The proposed array design facilitates a novel subspace-based DOA estimator, which allows estimating the DOAs of more uncorrelated sources than sensors, using only second-order statistics of the received data.
- A novel narrowband beamformer based on MD covariance fitting has been proposed, which provides an improved performance as compared to the state-of-the-art robust adaptive beamformers in scenarios with large sample support.
- Two broadband FIR beamformers based on worst-case output power minimization have been developed. These beamformers use different constraints to protect the desired signal component. It has been shown that the constraints of both beamformers strictly limit the sensitivity to signal steering vector estimation errors. Additionally, these constraints lead to an incentive for a low sensitivity if the true signal steering vectors lie within the presumed uncertainty sets. This incentive becomes stronger with increasing signal powers. The computational complexities of the proposed FIR beamformers are comparable to the complexity of the norm-bounded broadband MVDR beamformer.
- The relation between the proposed FIR beamformers and the norm-bounded broadband MVDR beamformer has been studied. Furthermore, the computationally efficient DFT beamformer counterparts of the proposed FIR beamformers have been formulated and analyzed.

- The proposed FIR beamformers control the frequency response towards the desired signal only for a discrete set of frequencies. To avoid the frequency discretization and associated errors, a modification of the first proposed FIR beamforming problem has been developed, which can be solved without frequency discretization. However, the modifications involve several approximations. Moreover, for large FIR filter lengths, the proposed FIR beamformer without frequency discretization has a higher complexity than the proposed FIR beamformers based on a frequency discretization.

- Simulation results show that the proposed FIR and DFT beamformers are attractive alternatives to the state-of-the-art robust adaptive broadband beamformers. In particular, the proposed FIR beamformers provide an improved performance as compared to previous FIR beamformers based on worst-case output power minimization.

The proposed concepts and methods have been developed from a general perspective without considering a particular application. Taking into account application specific properties of the system model, the proposed concepts and methods can be extended in many directions. For example, the signals transmitted in wireless communication systems are often finite alphabet or even constant modulus. The performance of the proposed beamformers can potentially be improved by exploiting this property.

For the proposed narrowband and broadband beamformers, we specified the orders of the computational complexities of the logarithmic barrier method. However, the orders describe the real computational complexity only approximately. Moreover, a number of modifications of and alternatives to the logarithmic barrier method exist, which have not been discussed. The computationally efficient implementation of the proposed beamformers including the development of tailored code is a crucial topic for real-time applications with many open questions.

The computational complexity of the proposed array geometry optimization problem is prohibitively high for arrays with large numbers of sensors. Therefore, another future research topic is the development of computationally efficient techniques to obtain good feasible points of the proposed array geometry optimization problem.

Appendix

A.1 Proof of Proposition 2.1

The L^2-norm of the approximation error of $a_n(\varphi)$ can be expressed as

$$\begin{aligned}\left\|a_n(\varphi) - S_{\frac{M_{\mathrm{MS}}-1}{2}}\{a_n\}(\varphi)\right\|_{L^2} &= \left(\int_{-\pi}^{\pi}\left|a_n(\varphi) - \sum_{m=-\frac{M_{\mathrm{MS}}-1}{2}}^{\frac{M_{\mathrm{MS}}-1}{2}} F_{a_n}[m]e^{jm\varphi}\right|^2 \mathrm{d}\varphi\right)^{1/2} \\ &= \left(\int_{-\pi}^{\pi}\left|\sum_{|m|>\frac{M_{\mathrm{MS}}-1}{2}} F_{a_n}[m]e^{jm\varphi}\right|^2 \mathrm{d}\varphi\right)^{1/2} \\ &= \left(2\pi \sum_{|m|>\frac{M_{\mathrm{MS}}-1}{2}} |F_{a_n}[m]|^2\right)^{1/2}.\end{aligned} \tag{A.1}$$

Hence, the L^2-norm of the approximation error depends on the magnitudes of the Fourier series coefficients, but not on their phases.

Let (p_{x_n}, p_{y_n}) and $(\tilde{p}_{x_n}, \tilde{p}_{y_n})$ have the same distance from the origin of the coordinate system. Hence,

$$\begin{bmatrix}\tilde{p}_{x_n} \\ \tilde{p}_{y_n}\end{bmatrix} = \begin{bmatrix}\cos\Delta\alpha & -\sin\Delta\alpha \\ \sin\Delta\alpha & \cos\Delta\alpha\end{bmatrix}\begin{bmatrix}p_{x_n} \\ p_{y_n}\end{bmatrix}, \tag{A.2}$$

where $\Delta\alpha$ is the angle between (p_{x_n}, p_{y_n}) and $(\tilde{p}_{x_n}, \tilde{p}_{y_n})$, seen from the origin of the coordinate system.

The mth Fourier series coefficient of

$$\tilde{a}_n(\varphi) = e^{j\frac{2\pi}{\lambda}(\tilde{p}_{x_n}\cos\varphi + \tilde{p}_{y_n}\sin\varphi)} \tag{A.3}$$

is given by

$$F_{\tilde{a}_n}[m] = \frac{1}{2\pi}\int_{-\pi}^{\pi}\tilde{a}_n(\varphi)\, e^{-jm\varphi}\, \mathrm{d}\varphi. \tag{A.4}$$

Since

$$\begin{aligned}\tilde{p}_{x_n}\cos\varphi + \tilde{p}_{y_n}\sin\varphi &= p_{x_n}\cos\Delta\alpha\cos\varphi - p_{y_n}\sin\Delta\alpha\cos\varphi \\ &\quad + p_{x_n}\sin\Delta\alpha\sin\varphi + p_{y_n}\cos\Delta\alpha\sin\varphi \\ &= p_{x_n}\cos(\varphi - \Delta\alpha) + p_{y_n}\sin(\varphi - \Delta\alpha),\end{aligned} \tag{A.5}$$

we obtain

$$\begin{aligned} F_{\tilde{a}_n}[m] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j\frac{2\pi}{\lambda}(p_{x_n}\cos(\varphi-\Delta\alpha)+p_{y_n}\sin(\varphi-\Delta\alpha))}\, e^{-jm\varphi}\,\mathrm{d}\varphi \\ &= \frac{1}{2\pi}\int_{-\pi-\Delta\alpha}^{\pi-\Delta\alpha} e^{j\frac{2\pi}{\lambda}(p_{x_n}\cos u+p_{y_n}\sin u)}\, e^{-jm(u+\Delta\alpha)}\,\mathrm{d}u \\ &= e^{-jm\Delta\alpha}F_{a_n}[m]. \end{aligned} \tag{A.6}$$

From (A.1) and (A.6), we obtain

$$\left\| a_n(\varphi) - S_{\frac{M_{\mathrm{MS}}-1}{2}}\{a_n\}(\varphi)\right\|_{L^2} = \left\| \tilde{a}_n(\varphi) - S_{\frac{M_{\mathrm{MS}}-1}{2}}\{\tilde{a}_n\}(\varphi)\right\|_{L^2}, \tag{A.7}$$

which was to be proved.

A.2 Proof of Proposition 2.2

Taking into account the conjugate reciprocity of the roots, the function $f_{\mathrm{MS},\mathsf{z}}(\mathsf{z})$ can be written as

$$f_{\mathrm{MS},\mathsf{z}}(\mathsf{z}) = \hat{c}\,\mathsf{z}^{M_{\mathrm{MS}}-1}\prod_{m=1}^{M_{\mathrm{MS}}-1}(1-\hat{\mathsf{z}}_m\mathsf{z}^{-1})(1-\hat{\mathsf{z}}_m^{*}\mathsf{z}), \tag{A.8}$$

where $\hat{\mathsf{z}}_m$ and $1/\hat{\mathsf{z}}_m^{*}$, $m = 1, \ldots, M_{\mathrm{MS}} - 1$ are the roots of $f_{\mathrm{MS},\mathsf{z}}(\mathsf{z})$, and the scalar $\hat{c}$ can be readily obtained using the equality of (2.34) and (A.8). The "hat" sign is used to indicate that $\hat{\mathsf{z}}_m$, $m = 1, \ldots, M_{\mathrm{MS}} - 1$ and $\hat{c}$ are functions of the sample covariance matrix $\hat{\boldsymbol{R}}_{\boldsymbol{x}}$.

If there are no manifold approximation and finite-sample errors, the function $f_{\mathrm{MS},\mathsf{z}}(\mathsf{z})$ will be equal to zero for $\mathsf{z}_l = e^{j\varphi_l}$, $l = 1, \ldots, L$. The manifold approximation and finite-sample errors, however, lead to the roots (see Figure A.1)

$$\hat{\mathsf{z}}_l = \mathsf{z}_l + \Delta\mathsf{z}_l = |\hat{\mathsf{z}}_l|e^{j\hat{\varphi}_l}, \quad l = 1, \ldots, L. \tag{A.9}$$

From Figure A.1, we obtain

$$\frac{\sin(\hat{\varphi}_l - \varphi_l)}{r_l} = \frac{\sin(2\pi - (\pi - \varphi_l) - \Phi_l)}{|\hat{\mathsf{z}}_l|}, \tag{A.10}$$

where r_l and Φ_l are the magnitude and phase of

$$\Delta\mathsf{z}_l = r_l e^{j\Phi_l}, \tag{A.11}$$

respectively. If $r_l \ll 1$, (A.10) yields

$$\hat{\varphi}_l - \varphi_l = \arcsin\left(r_l\frac{\sin(\Phi_l - \varphi_l)}{|\hat{\mathsf{z}}_l|}\right) \approx r_l\sin(\Phi_l - \varphi_l). \tag{A.12}$$

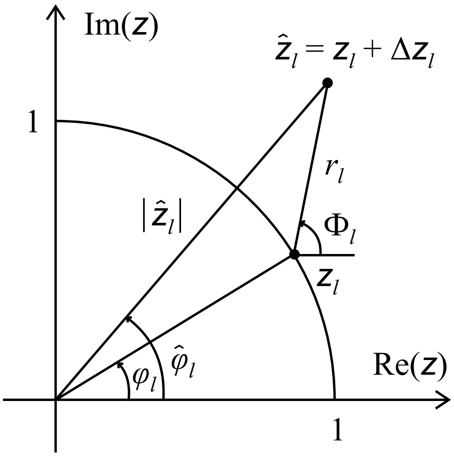


Figure A.1. Root location.

From (2.32), (2.34), and (A.8), we obtain

$$f_{\mathrm{MS}}(\varphi) = \hat{c} \prod_{m=1}^{M_{\mathrm{MS}}-1} (1 - \hat{\mathsf{z}}_m e^{-j\varphi})(1 - \hat{\mathsf{z}}_m^* e^{j\varphi}) = \hat{c} \prod_{m=1}^{M_{\mathrm{MS}}-1} \hat{p}_m(\varphi), \tag{A.13}$$

where

$$\hat{p}_m(\varphi) = (1 - \hat{\mathsf{z}}_m e^{-j\varphi})(1 - \hat{\mathsf{z}}_m^* e^{j\varphi}). \tag{A.14}$$

Using (A.9) and (A.11), $\hat{p}_l(\varphi_l)$, $l = 1, \ldots, L$, and its first and second derivatives can be expressed as

$$\hat{p}_l(\varphi_l) = r_l^2 \tag{A.15}$$

$$\hat{p}_l'(\varphi_l) = 2\, r_l \sin(\varphi_l - \Phi_l) \tag{A.16}$$

$$\hat{p}_l''(\varphi_l) = 2 + 2\, r_l \cos(\varphi_l - \Phi_l) \approx 2. \tag{A.17}$$

Using (A.15) and (A.16), we obtain for the derivative of $f_{\mathrm{MS}}(\varphi_l)$ that

$$\begin{aligned} f_{\mathrm{MS}}'(\varphi_l) &= 2\,\hat{c}\, r_l \,\sin(\varphi_l - \Phi_l) \prod_{\substack{m=1\\ m\neq l}}^{M_{\mathrm{MS}}-1} \hat{p}_m(\varphi_l) + \hat{c}\, r_l^2 \sum_{\substack{m=1\\ m\neq l}}^{M_{\mathrm{MS}}-1} \hat{p}_m'(\varphi_l) \prod_{\substack{k=1\\ k\neq l,m}}^{M_{\mathrm{MS}}-1} \hat{p}_k(\varphi_l) \\ &\approx 2\,\hat{c}\, r_l \,\sin(\varphi_l - \Phi_l) \prod_{\substack{m=1\\ m\neq l}}^{M_{\mathrm{MS}}-1} \hat{p}_m(\varphi_l). \end{aligned} \tag{A.18}$$

Therefore, (A.12) can be reformulated as

$$\hat{\varphi}_l - \varphi_l \approx -\frac{f_{\mathrm{MS}}'(\varphi_l)}{2\,\hat{c} \prod\limits_{\substack{m=1\\ m\neq l}}^{M_{\mathrm{MS}}-1} \hat{p}_m(\varphi_l)}. \tag{A.19}$$

Similarly, from (A.15)-(A.17) we obtain that the second derivative of $f_{\mathrm{MS}}(\varphi)$ at $\varphi = \varphi_l$ is given by

$$f''_{\mathrm{MS}}(\varphi_l) \approx 2\,\hat{c} \prod_{\substack{m=1\\ m\neq l}}^{M_{\mathrm{MS}}-1} \hat{p}_m(\varphi_l). \tag{A.20}$$

Inserting (A.20) into (A.19), we obtain

$$\hat{\varphi}_l - \varphi_l \approx -\frac{f'_{\mathrm{MS}}(\varphi_l)}{f''_{\mathrm{MS}}(\varphi_l)}. \tag{A.21}$$

A first-order approximation of the second derivative of $f_{\mathrm{MS}}(\varphi)$ at $\varphi = \varphi_l$ yields

$$\begin{aligned} f''_{\mathrm{MS}}(\varphi_l) = {} & \frac{\partial^2 \boldsymbol{g}^H_{\mathrm{MS}}(\varphi_l)}{\partial\varphi^2}\, \boldsymbol{G}^H_{\mathrm{MS}}\, \hat{\boldsymbol{V}}_{\boldsymbol{x},N} \hat{\boldsymbol{V}}^H_{\boldsymbol{x},N}\, \boldsymbol{G}_{\mathrm{MS}}\, \boldsymbol{g}_{\mathrm{MS}}(\varphi_l) \\ & + \boldsymbol{g}^H_{\mathrm{MS}}(\varphi_l)\, \boldsymbol{G}^H_{\mathrm{MS}}\, \hat{\boldsymbol{V}}_{\boldsymbol{x},N} \hat{\boldsymbol{V}}^H_{\boldsymbol{x},N}\, \boldsymbol{G}_{\mathrm{MS}}\, \frac{\partial^2 \boldsymbol{g}_{\mathrm{MS}}(\varphi_l)}{\partial\varphi^2} \\ & + 2\frac{\partial \boldsymbol{g}^H_{\mathrm{MS}}(\varphi_l)}{\partial\varphi}\, \boldsymbol{G}^H_{\mathrm{MS}}\, \hat{\boldsymbol{V}}_{\boldsymbol{x},N} \hat{\boldsymbol{V}}^H_{\boldsymbol{x},N}\, \boldsymbol{G}_{\mathrm{MS}}\, \frac{\partial \boldsymbol{g}_{\mathrm{MS}}(\varphi_l)}{\partial\varphi} \\ \approx {} & 2\frac{\partial \boldsymbol{g}^H_{\mathrm{MS}}(\varphi_l)}{\partial\varphi}\, \boldsymbol{G}^H_{\mathrm{MS}}\, \hat{\boldsymbol{V}}_{\boldsymbol{x},N} \hat{\boldsymbol{V}}^H_{\boldsymbol{x},N}\, \boldsymbol{G}_{\mathrm{MS}}\, \frac{\partial \boldsymbol{g}_{\mathrm{MS}}(\varphi_l)}{\partial\varphi} \\ \approx {} & 2\frac{\partial \boldsymbol{g}^H_{\mathrm{MS}}(\varphi_l)}{\partial\varphi}\, \boldsymbol{G}^H_{\mathrm{MS}}\, \boldsymbol{V}_{\boldsymbol{x},N} \boldsymbol{V}^H_{\boldsymbol{x},N}\, \boldsymbol{G}_{\mathrm{MS}}\, \frac{\partial \boldsymbol{g}_{\mathrm{MS}}(\varphi_l)}{\partial\varphi}. \end{aligned} \tag{A.22}$$

Equations (A.21) and (A.22) give

$$\mathrm{E}\{(\hat{\varphi}_l - \varphi_l)^2\} \approx \frac{\mathrm{E}\{(f'_{\mathrm{MS}}(\varphi_l))^2\}}{\left(2\frac{\partial \boldsymbol{g}^H_{\mathrm{MS}}(\varphi_l)}{\partial\varphi}\, \boldsymbol{G}^H_{\mathrm{MS}}\, \boldsymbol{V}_{\boldsymbol{x},N} \boldsymbol{V}^H_{\boldsymbol{x},N}\, \boldsymbol{G}_{\mathrm{MS}}\, \frac{\partial \boldsymbol{g}_{\mathrm{MS}}(\varphi_l)}{\partial\varphi}\right)^2}. \tag{A.23}$$

Let us introduce the eigenvector estimation error

$$\boldsymbol{\delta}_{\boldsymbol{x},k} = \hat{\boldsymbol{v}}_{\boldsymbol{x},k} - \boldsymbol{v}_{\boldsymbol{x},k}, \quad k = 1, \ldots, N. \tag{A.24}$$

The following relations will be used [KB86b, KB86a]:

$$\mathrm{E}\{\boldsymbol{\delta}_{\boldsymbol{x},k}\boldsymbol{\delta}^H_{\boldsymbol{x},m}\} = \frac{\lambda_{\boldsymbol{x},k}}{K} \sum_{\substack{n=1\\ n\neq k}}^{N} \frac{\lambda_{\boldsymbol{x},n}}{(\lambda_{\boldsymbol{x},k} - \lambda_{\boldsymbol{x},n})^2}\, \boldsymbol{v}_{\boldsymbol{x},n}\, \boldsymbol{v}^H_{\boldsymbol{x},n}\, \delta_{k,m} + o(K^{-1}), \quad 1 \le k, m \le L \tag{A.25}$$

$$\mathrm{E}\{\boldsymbol{\delta}_{\boldsymbol{x},k}\boldsymbol{\delta}^T_{\boldsymbol{x},m}\} = -\frac{\lambda_{\boldsymbol{x},k}\lambda_{\boldsymbol{x},m}}{K(\lambda_{\boldsymbol{x},k} - \lambda_{\boldsymbol{x},m})^2}\, \boldsymbol{v}_{\boldsymbol{x},m}\, \boldsymbol{v}^T_{\boldsymbol{x},k}(1 - \delta_{k,m}) + o(K^{-1}), \quad 1 \le k, m \le L \tag{A.26}$$

$$\mathrm{E}\{\boldsymbol{\delta}_{\boldsymbol{x},k}\} = -\frac{\lambda_{\boldsymbol{x},k}}{2K} \sum_{\substack{n=1\\ n\neq k}}^{N} \frac{\lambda_{\boldsymbol{x},n}}{(\lambda_{\boldsymbol{x},k} - \lambda_{\boldsymbol{x},n})^2}\, \boldsymbol{v}_{\boldsymbol{x},k} + o(K^{-1}), \quad 1 \le k \le L, \tag{A.27}$$

where $\delta_{k,m}$ denotes the Kronecker delta.

From (2.48), we have

$$f'_{\mathrm{MS}}(\varphi_l) = D_{1,l} + D^*_{1,l} - D_{2,l} - D^*_{2,l} - D_{3,l} - D^*_{3,l}, \tag{A.28}$$

where

$$D_{1,l} = \frac{\partial \boldsymbol{g}^H_{\mathrm{MS}}(\varphi_l)}{\partial \varphi}\, \boldsymbol{G}^H_{\mathrm{MS}}\, \boldsymbol{V}_{\boldsymbol{x},N} \boldsymbol{V}^H_{\boldsymbol{x},N}\, \boldsymbol{G}_{\mathrm{MS}}\, \boldsymbol{g}_{\mathrm{MS}}(\varphi_l) \tag{A.29}$$

$$D_{2,l} = \sum_{k=1}^{L} \frac{\partial \boldsymbol{g}^H_{\mathrm{MS}}(\varphi_l)}{\partial \varphi}\, \boldsymbol{G}^H_{\mathrm{MS}}\, (\boldsymbol{v}_{\boldsymbol{x},k}\, \boldsymbol{\delta}^H_{\boldsymbol{x},k} + \boldsymbol{\delta}_{\boldsymbol{x},k}\, \boldsymbol{v}^H_{\boldsymbol{x},k})\, \boldsymbol{G}_{\mathrm{MS}}\, \boldsymbol{g}_{\mathrm{MS}}(\varphi_l) \tag{A.30}$$

$$D_{3,l} = \sum_{k=1}^{L} \frac{\partial \boldsymbol{g}^H_{\mathrm{MS}}(\varphi_l)}{\partial \varphi}\, \boldsymbol{G}^H_{\mathrm{MS}}\, \boldsymbol{\delta}_{\boldsymbol{x},k}\, \boldsymbol{\delta}^H_{\boldsymbol{x},k}\, \boldsymbol{G}_{\mathrm{MS}}\, \boldsymbol{g}_{\mathrm{MS}}(\varphi_l). \tag{A.31}$$

Hence, the numerator of (A.23) is given by

$$\begin{aligned} \mathrm{E}\{(f'_{\mathrm{MS}}(\varphi_l))^2\} = {} & \mathrm{E}\{(D_{1,l} + D^*_{1,l})^2\} + \mathrm{E}\{(D_{2,l} + D^*_{2,l})^2\} + \mathrm{E}\{(D_{3,l} + D^*_{3,l})^2\} \\ & - 2\mathrm{E}\{(D_{1,l} + D^*_{1,l})(D_{2,l} + D^*_{2,l})\} \\ & - 2\mathrm{E}\{(D_{1,l} + D^*_{1,l})(D_{3,l} + D^*_{3,l})\} \\ & - 2\mathrm{E}\{(D_{2,l} + D^*_{2,l})(D_{3,l} + D^*_{3,l})\}. \end{aligned} \tag{A.32}$$

Neglecting all terms that contain higher than the second orders of $\boldsymbol{\delta}_{\boldsymbol{x},k}$, we obtain

$$\begin{aligned} \mathrm{E}\{(f'_{\mathrm{MS}}(\varphi_l))^2\} \approx {} & \mathrm{E}\{(D_{1,l} + D^*_{1,l})^2\} + \mathrm{E}\{(D_{2,l} + D^*_{2,l})^2\} \\ & - 2\mathrm{E}\{(D_{1,l} + D^*_{1,l})(D_{2,l} + D^*_{2,l})\} \\ & - 2\mathrm{E}\{(D_{1,l} + D^*_{1,l})(D_{3,l} + D^*_{3,l})\}. \end{aligned} \tag{A.33}$$

As $D_{1,l}$ is constant, (A.33) can be simplified as

$$\begin{aligned} \mathrm{E}\{(f'_{\mathrm{MS}}(\varphi_l))^2\} \approx {} & 4\Re^2\{D_{1,l}\} + 2\Re\{\mathrm{E}\{D^2_{2,l}\}\} + 2\mathrm{E}\{|D_{2,l}|^2\} \\ & - 8\Re\{D_{1,l}\}\Re\{\mathrm{E}\{D_{2,l}\}\} - 8\Re\{D_{1,l}\}\Re\{\mathrm{E}\{D_{3,l}\}\}. \end{aligned} \tag{A.34}$$

Using (A.25)–(A.27), we have

$$\mathrm{E}\{D^2_{2,l}\} \approx \sum_{k=1}^{L} \sum_{n=L+1}^{N} 2\mu_{k,n}\, \boldsymbol{d}^H_{\mathrm{MS},l} \boldsymbol{v}_{\boldsymbol{x},k} \boldsymbol{d}^H_{\mathrm{MS},l} \boldsymbol{v}_{\boldsymbol{x},n} \boldsymbol{a}^T_{\mathrm{MS},l} \boldsymbol{v}^*_{\boldsymbol{x},k} \boldsymbol{a}^T_{\mathrm{MS},l} \boldsymbol{v}^*_{\boldsymbol{x},n} \tag{A.35}$$

$$\mathrm{E}\{|D_{2,l}|^2\} \approx \sum_{k=1}^{L} \sum_{n=L+1}^{N} \mu_{k,n} \big(|\boldsymbol{d}^H_{\mathrm{MS},l} \boldsymbol{v}_{\boldsymbol{x},k}|^2 |\boldsymbol{a}^H_{\mathrm{MS},l} \boldsymbol{v}_{\boldsymbol{x},n}|^2 + |\boldsymbol{a}^H_{\mathrm{MS},l} \boldsymbol{v}_{\boldsymbol{x},k}|^2 |\boldsymbol{d}^H_{\mathrm{MS},l} \boldsymbol{v}_{\boldsymbol{x},n}|^2 \big) \tag{A.36}$$

$$\mathrm{E}\{D_{2,l}\} \approx -\sum_{k=1}^{L} \sum_{\substack{n=1 \\ n \neq k}}^{N} \mu_{k,n} \boldsymbol{d}^H_{\mathrm{MS},l} \boldsymbol{v}_{\boldsymbol{x},k} \boldsymbol{v}^H_{\boldsymbol{x},k} \boldsymbol{a}_{\mathrm{MS},l} \tag{A.37}$$

$$\mathrm{E}\{D_{3,l}\} \approx \sum_{k=1}^{L} \sum_{\substack{n=1 \\ n \neq k}}^{N} \mu_{k,n} \boldsymbol{d}^H_{\mathrm{MS},l} \boldsymbol{v}_{\boldsymbol{x},n} \boldsymbol{v}^H_{\boldsymbol{x},n} \boldsymbol{a}_{\mathrm{MS},l}. \tag{A.38}$$

Inserting (A.35)–(A.38) into (A.34) and then substituting (A.34) into (A.23), we prove (2.41).

A.3 Proof of Proposition 4.2

The objective and the first constraint functions of (4.49) are invariant to the multiplication of $\breve{\boldsymbol{A}}$ with an arbitrary unitary matrix. The constraints in the third and fourth lines of (4.49) are not subject to this invariance. However, the constraints in the third line of (4.49) may be inactive. This occurs, for example, if

$$\hat{\boldsymbol{R}}_{\boldsymbol{x}} = \boldsymbol{R}_{\boldsymbol{x}}, \ \hat{\boldsymbol{a}}_l = \boldsymbol{a}_l, \text{ and } \epsilon_l > 0, \quad l = 1, \ldots, L. \tag{A.39}$$

Then, $(\breve{\boldsymbol{A}}_{\mathrm{MD}}, \sigma_{\mathrm{MD}}) = (\boldsymbol{A}\boldsymbol{P}^{1/2}, \sigma)$ is a solution of (4.49) with inactive constraints in the third line of (4.49).

Let $(\breve{\boldsymbol{A}}_{\mathrm{MD}}, \sigma_{\mathrm{MD}})$ be an arbitrary solution of (4.49). Then,

$$\Re\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}}_{\mathrm{MD}} \, (\boldsymbol{e}_l + \breve{\boldsymbol{\delta}}_l)\} \geq \Re\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}}_{\mathrm{MD}} \, \boldsymbol{e}_l\} - \|\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}}_{\mathrm{MD}}\|_2 \, \|\breve{\boldsymbol{\delta}}_l\|_2, \tag{A.40}$$

and

$$\nu_l \|\breve{\boldsymbol{A}}_{\mathrm{MD}} \, (\boldsymbol{e}_l + \breve{\boldsymbol{\delta}}_l)\|_2 \leq \nu_l \|\breve{\boldsymbol{A}}_{\mathrm{MD}} \, \boldsymbol{e}_l\|_2 + \nu_l \|\breve{\boldsymbol{A}}_{\mathrm{MD}}\|_2 \, \|\breve{\boldsymbol{\delta}}_l\|_2. \tag{A.41}$$

Hence,

$$\Re\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}}_{\mathrm{MD}} \, (\boldsymbol{e}_l + \breve{\boldsymbol{\delta}}_l)\} \geq \nu_l \, \|\breve{\boldsymbol{A}}_{\mathrm{MD}} \, (\boldsymbol{e}_l + \breve{\boldsymbol{\delta}}_l)\|_2 \tag{A.42}$$

if

$$\|\breve{\boldsymbol{\delta}}_l\|_2 \leq \frac{\Re\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}}_{\mathrm{MD}} \, \boldsymbol{e}_l\} - \nu_l \, \|\breve{\boldsymbol{A}}_{\mathrm{MD}} \, \boldsymbol{e}_l\|_2}{\|\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}}_{\mathrm{MD}}\|_2 + \nu_l \, \|\breve{\boldsymbol{A}}_{\mathrm{MD}}\|_2} \triangleq \xi_l. \tag{A.43}$$

If the constraints in the third line of (4.49) are inactive for $\breve{\boldsymbol{A}}_{\mathrm{MD}}$, then $\xi_l > 0$ for $l = 1, \ldots, L$. Let $\xi_{\min}$ denote the minimum of $\xi_1, \ldots, \xi_L$, and

$$\boldsymbol{U} \triangleq \boldsymbol{I}_L + \left[\breve{\boldsymbol{\delta}}_1, \ldots, \breve{\boldsymbol{\delta}}_L\right]. \tag{A.44}$$

If the $L \times L$ matrix $\boldsymbol{U}$ satisfies

$$\|\boldsymbol{U} - \boldsymbol{I}_L\|_F \leq \xi_{\min}, \tag{A.45}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, then $\breve{\boldsymbol{A}}_{\mathrm{MD}}\boldsymbol{U}$ satisfies the constraints in the third line of (4.49). Note that the set of unitary $L \times L$ matrices forms a smooth manifold of dimension L^2 [Hal03]. Hence, the set of unitary $L \times L$ matrices $\boldsymbol{U}$ with $\|\boldsymbol{U} - \boldsymbol{I}_L\|_F \leq \xi_{\min}$ is nonempty. For any such matrix, we have that

$$\begin{aligned} \hat{\boldsymbol{a}}_l \breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U} \boldsymbol{D}(\hat{\boldsymbol{A}}, \breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U}) \boldsymbol{e}_l &= \Re\{\hat{\boldsymbol{a}}_l \breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U} \boldsymbol{D}(\hat{\boldsymbol{A}}, \breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U}) \boldsymbol{e}_l\} \\ &\geq \Re\{\hat{\boldsymbol{a}}_l \breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U} \boldsymbol{e}_l\} \\ &\geq \nu_l \, \|\breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U} \boldsymbol{e}_l\|_2 \\ &= \nu_l \, \|\breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U} \boldsymbol{D}(\hat{\boldsymbol{A}}, \breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U}) \boldsymbol{e}_l\|_2, \end{aligned} \tag{A.46}$$

where $\boldsymbol{D}(\hat{\boldsymbol{A}}, \breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U})$ is defined in (4.54). Hence, the constraints in the third and fourth lines of (4.49) are satisfied for all $\breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U} \boldsymbol{D}(\hat{\boldsymbol{A}}, \breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U})$. Consequently, all $(\breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U} \boldsymbol{D}(\hat{\boldsymbol{A}}, \breve{\boldsymbol{A}}_{\mathrm{MD}} \boldsymbol{U}), \sigma_{\mathrm{MD}})$ are optimum points, i.e., the MD covariance fitting problem does not have a unique optimum point in general.

A.4 Proof of Proposition 4.3

The proof of this proposition follows immediately from the proof of Proposition 4.2, since

$$\boldsymbol{R}_{\boldsymbol{x}} = \breve{\boldsymbol{A}} \breve{\boldsymbol{A}}^H + \sigma \boldsymbol{I}_N \tag{A.47}$$

for all

$$\breve{\boldsymbol{A}} = \boldsymbol{A} \boldsymbol{P}^{1/2} \boldsymbol{U} \boldsymbol{D}(\boldsymbol{A}, \boldsymbol{A} \boldsymbol{P}^{1/2} \boldsymbol{U}), \tag{A.48}$$

where $\boldsymbol{U}$ is an arbitrary unitary $L \times L$ matrix that satisfies $\|\boldsymbol{U} - \boldsymbol{I}_L\|_F \leq \xi_{\min}$,

$$\xi_{\min} = \min_l \frac{\Re\{\boldsymbol{a}_l^H \boldsymbol{A} \boldsymbol{P}^{1/2} \boldsymbol{e}_l\} - \sqrt{\|\boldsymbol{a}_l\|_2^2 - \epsilon_l^2} \, \|\boldsymbol{A} \boldsymbol{P}^{1/2} \boldsymbol{e}_l\|_2}{\|\boldsymbol{a}_l^H \boldsymbol{A} \boldsymbol{P}^{1/2}\|_2 + \sqrt{\|\boldsymbol{a}_l\|_2^2 - \epsilon_l^2} \, \|\boldsymbol{A} \boldsymbol{P}^{1/2}\|_2}, \tag{A.49}$$

and $\boldsymbol{D}(\boldsymbol{A}, \boldsymbol{A} \boldsymbol{P}^{1/2} \boldsymbol{U})$ is defined in (4.54).

A.5 Proof of Proposition 4.4

We first show that the signal steering vector is essentially identifiable under the conditions of Proposition 4.4. Thereto, we show that the constraints

$$\begin{aligned} &\breve{\boldsymbol{A}} \breve{\boldsymbol{A}}^H + \tilde{\sigma} \boldsymbol{I}_N = \boldsymbol{R}_{\boldsymbol{x}} \\ &\breve{\boldsymbol{A}} \boldsymbol{e}_l \doteq \boldsymbol{A} \boldsymbol{e}_l, \quad l = 2, \ldots, L, \end{aligned} \tag{A.50}$$

imply that $\breve{\boldsymbol{A}} \boldsymbol{e}_1 \doteq \boldsymbol{A} \boldsymbol{e}_1$. Since $L < N$, the smallest eigenvalue of $\boldsymbol{R}_{\boldsymbol{x}}$ is equal to $\sigma = \tilde{\sigma}$. Hence, $\breve{\boldsymbol{A}}$ has to satisfy

$$\breve{\boldsymbol{A}} \breve{\boldsymbol{A}}^H = \boldsymbol{A} \boldsymbol{P} \boldsymbol{A}^H. \tag{A.51}$$

Equation (A.51) gives

$$\breve{\boldsymbol{A}} = \boldsymbol{A}\boldsymbol{B}, \tag{A.52}$$

where $\boldsymbol{B}$ is an $L \times L$ matrix. Due to the constraints in the second line of (A.50) and the linear independence of the source steering vectors, we have

$$\boldsymbol{B} = \boldsymbol{b}\boldsymbol{e}_1^T + \operatorname{diag}(0, \beta_2, \ldots \beta_L), \tag{A.53}$$

where $\boldsymbol{b}$ is an $L \times 1$ vector. Substituting (A.52) in (A.51) yields

$$\boldsymbol{P} = \boldsymbol{B}\boldsymbol{B}^H, \tag{A.54}$$

where we have used again that the source steering vectors are linearly independent. Since

$$\boldsymbol{B}\boldsymbol{B}^H = \boldsymbol{b}\boldsymbol{b}^H + \operatorname{diag}(0, |\beta_2|^2, \ldots |\beta_L|^2), \tag{A.55}$$

we have that $\boldsymbol{b}$ is a scalar multiple of $\boldsymbol{e}_1$. Hence, $\breve{\boldsymbol{A}}\boldsymbol{e}_1 \doteq \boldsymbol{A}\boldsymbol{e}_1$, so the signal steering vector is essentially identifiable.

Let us now show that

$$(\breve{\boldsymbol{A}}_{\mathrm{MD}}, \sigma_{\mathrm{MD}}) = (\boldsymbol{A}\boldsymbol{P}^{1/2}\boldsymbol{D}(\hat{\boldsymbol{A}}, \boldsymbol{A}), \sigma) \tag{A.56}$$

is the unique optimum point of (4.49) under the conditions of Proposition 4.4. Note that the objective function for $(\breve{\boldsymbol{A}}_{\mathrm{MD}}, \sigma_{\mathrm{MD}})$ is $\log\det \boldsymbol{R}_{\boldsymbol{x}}$. It directly follows from the constraint in the second line of (4.49) that $(\breve{\boldsymbol{A}}, \tilde{\sigma})$ can only be an optimum point if

$$\breve{\boldsymbol{A}}\breve{\boldsymbol{A}}^H + \tilde{\sigma}\boldsymbol{I}_N = \boldsymbol{R}_{\boldsymbol{x}}. \tag{A.57}$$

Hence, the columns of $\breve{\boldsymbol{A}}$ have to be essentially equal to the true source steering vectors. We have shown in Section 4.2.2 that the columns of $\breve{\boldsymbol{A}}$ are not subject to any scaling ambiguity, so (A.56) is the unique optimum point of (4.49) under the conditions of Proposition 4.4.

A.6 Computational complexity of (4.67)

The barrier function corresponding to (4.67) can be expressed as [BV04]

$$\begin{aligned} b(\breve{\boldsymbol{A}}, \tilde{\sigma}, \boldsymbol{\kappa}, t) = & -2\Re\left\{\operatorname{tr}\left\{\breve{\boldsymbol{A}}_0^H \boldsymbol{R}_{\boldsymbol{x},0}^{-1}\breve{\boldsymbol{A}}\right\}\right\} - \operatorname{tr}\left\{\boldsymbol{R}_{\boldsymbol{x},0}^{-1}\right\}\tilde{\sigma} - \frac{1}{t}\log\det \boldsymbol{F}(\breve{\boldsymbol{A}}, \tilde{\sigma}) \\ & - \frac{1}{t}\sum_{l=1}^{L}\left(\log\left(\Re\{\hat{\boldsymbol{a}}_l^H\breve{\boldsymbol{A}}\boldsymbol{e}_l\} - \kappa_l\right) + \log\left(\kappa_l^2 - \nu_l^2\left\|\breve{\boldsymbol{A}}\boldsymbol{e}_l\right\|_2^2\right)\right) \\ & - \frac{1}{t}\sum_{l=1}^{L}\log\kappa_l - \frac{1}{t}\log\tilde{\sigma}, \end{aligned} \tag{A.58}$$

where $\boldsymbol{\kappa} = [\kappa_1, \ldots, \kappa_L]^T$ is an $L \times 1$ vector of positive slack variables, and $t > 0$. If $t \to \infty$, then the logarithmic barrier functions converge to indicator functions [BV04]. Hence, the optimum points of

$$\begin{aligned} \min_{\breve{\boldsymbol{A}}, \tilde{\sigma}, \boldsymbol{\kappa}} \quad & b(\breve{\boldsymbol{A}}, \tilde{\sigma}, \boldsymbol{\kappa}, t) \\ \text{s.t.} \quad & \Im\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{A}} \boldsymbol{e}_l\} = 0, \quad l = 1, \ldots, L \end{aligned} \tag{A.59}$$

converge to optimum points of (4.67) if $t \to \infty$.

To solve (A.59), we use Newton's method [BV04]. Let

$$\boldsymbol{v} = \begin{bmatrix} \text{vec}\{\Re\{\breve{\boldsymbol{A}}\}\} \\ \text{vec}\{\Im\{\breve{\boldsymbol{A}}\}\} \\ \tilde{\sigma} \\ \boldsymbol{\kappa} \end{bmatrix}, \tag{A.60}$$

denote the $(2LN + L + 1) \times 1$ real-valued vector of optimization variables, where vec$\{\cdot\}$ is the vectorization operator. Then, the quadratic Taylor approximation of the barrier function about $\boldsymbol{v}$ can be written as

$$b(\boldsymbol{v} + \Delta\boldsymbol{v}, t) \approx b(\boldsymbol{v}, t) + \Delta\boldsymbol{v}^T \frac{\partial b(\boldsymbol{v}, t)}{\partial \boldsymbol{v}} + \frac{1}{2} \Delta\boldsymbol{v}^T \frac{\partial^2 b(\boldsymbol{v}, t)}{\partial \boldsymbol{v} \partial \boldsymbol{v}^T} \Delta\boldsymbol{v}. \tag{A.61}$$

Moreover, the constraints in (A.59) can be expressed as

$$\boldsymbol{G}\boldsymbol{v} = \boldsymbol{0}, \tag{A.62}$$

where $\boldsymbol{G}$ is an $L \times (2LN + L + 1)$ matrix. The lth row of this matrix is given by

$$\boldsymbol{G}_{l,:} = \begin{bmatrix} -\boldsymbol{e}_l \otimes \Im\{\hat{\boldsymbol{a}}_l\} \\ \boldsymbol{e}_l \otimes \Re\{\hat{\boldsymbol{a}}_l\} \\ \boldsymbol{0}_{L+1} \end{bmatrix}^T, \tag{A.63}$$

where $\otimes$ denotes the Kronecker matrix product. Newton's method involves solving a sequence of quadratic approximations of (A.59)

$$\begin{aligned} \min_{\Delta\boldsymbol{v}} \quad & b(\boldsymbol{v}, t) + \Delta\boldsymbol{v}^T \tfrac{\partial b(\boldsymbol{v}, t)}{\partial \boldsymbol{v}} + \tfrac{1}{2} \Delta\boldsymbol{v}^T \tfrac{\partial^2 b(\boldsymbol{v}, t)}{\partial \boldsymbol{v} \partial \boldsymbol{v}^T} \Delta\boldsymbol{v} \\ \text{s.t.} \quad & \boldsymbol{G}\boldsymbol{v} = \boldsymbol{0}. \end{aligned} \tag{A.64}$$

The KKT system corresponding to (A.64) can be expressed as

$$\begin{bmatrix} \frac{\partial^2 b}{\partial \boldsymbol{v} \partial \boldsymbol{v}^T} & \boldsymbol{G}^T \\ \boldsymbol{G} & \boldsymbol{0}_{L \times L} \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{v} \\ \boldsymbol{u} \end{bmatrix} = \begin{bmatrix} -\frac{\partial b}{\partial \boldsymbol{v}} \\ \boldsymbol{0}_{L \times 1} \end{bmatrix}, \tag{A.65}$$

where $\boldsymbol{u}$ is an $L \times 1$ vector of Lagrange parameters. Any vector $\left[\Delta\boldsymbol{v}^T, \boldsymbol{u}^T\right]^T$, which satisfies (A.65), is an optimum point of (A.64). The gradient and Hessian of the barrier function can be written as

$$\frac{\partial b}{\partial \boldsymbol{v}} = \begin{bmatrix} \frac{\partial b}{\partial \mathrm{vec}\{\Re\{\breve{\boldsymbol{A}}\}\}} \\ \frac{\partial b}{\partial \mathrm{vec}\{\Im\{\breve{\boldsymbol{A}}\}\}} \\ \frac{\partial b}{\partial \tilde{\sigma}} \\ \frac{\partial b}{\partial \boldsymbol{\kappa}} \end{bmatrix} \tag{A.66}$$

and

$$\frac{\partial^2 b}{\partial \boldsymbol{v} \partial \boldsymbol{v}^T} = \begin{bmatrix} \frac{\partial^2 b}{\partial \mathrm{vec}\{\Re\{\breve{\boldsymbol{A}}\}\}\partial \mathrm{vec}^T\{\Re\{\breve{\boldsymbol{A}}\}\}} & \frac{\partial^2 b}{\partial \mathrm{vec}\{\Re\{\breve{\boldsymbol{A}}\}\}\partial \mathrm{vec}^T\{\Im\{\breve{\boldsymbol{A}}\}\}} & \frac{\partial^2 b}{\partial \mathrm{vec}\{\Re\{\breve{\boldsymbol{A}}\}\}\partial \tilde{\sigma}} & \frac{\partial^2 b}{\partial \mathrm{vec}\{\Re\{\breve{\boldsymbol{A}}\}\}\partial \boldsymbol{\kappa}^T} \\ \frac{\partial^2 b}{\partial \mathrm{vec}\{\Im\{\breve{\boldsymbol{A}}\}\}\partial \mathrm{vec}^T\{\Re\{\breve{\boldsymbol{A}}\}\}} & \frac{\partial^2 b}{\partial \mathrm{vec}\{\Im\{\breve{\boldsymbol{A}}\}\}\partial \mathrm{vec}^T\{\Im\{\breve{\boldsymbol{A}}\}\}} & \frac{\partial^2 b}{\partial \mathrm{vec}\{\Im\{\breve{\boldsymbol{A}}\}\}\partial \tilde{\sigma}} & \frac{\partial^2 b}{\partial \mathrm{vec}\{\Im\{\breve{\boldsymbol{A}}\}\}\partial \boldsymbol{\kappa}^T} \\ \frac{\partial^2 b}{\partial \tilde{\sigma} \partial \mathrm{vec}^T\{\Re\{\breve{\boldsymbol{A}}\}\}} & \frac{\partial^2 b}{\partial \tilde{\sigma} \partial \mathrm{vec}^T\{\Im\{\breve{\boldsymbol{A}}\}\}} & \frac{\partial^2 b}{\partial \tilde{\sigma}^2} & \frac{\partial^2 b}{\partial \tilde{\sigma} \partial \boldsymbol{\kappa}^T} \\ \frac{\partial^2 b}{\partial \boldsymbol{\kappa} \partial \mathrm{vec}^T\{\Re\{\breve{\boldsymbol{A}}\}\}} & \frac{\partial^2 b}{\partial \boldsymbol{\kappa} \partial \mathrm{vec}^T\{\Im\{\breve{\boldsymbol{A}}\}\}} & \frac{\partial^2 b}{\partial \boldsymbol{\kappa} \partial \tilde{\sigma}} & \frac{\partial^2 b}{\partial \boldsymbol{\kappa} \partial \boldsymbol{\kappa}^T} \end{bmatrix}, \tag{A.67}$$

respectively.

To compute the gradient (A.66), we use [Bra83]

$$\frac{\partial b}{\partial \Re\{\breve{\boldsymbol{A}}\}} = 2\Re\left\{\frac{\partial b}{\partial \breve{\boldsymbol{A}}^*}\right\} \qquad \frac{\partial b}{\partial \Im\{\breve{\boldsymbol{A}}\}} = 2\Im\left\{\frac{\partial b}{\partial \breve{\boldsymbol{A}}^*}\right\}. \tag{A.68}$$

Equation (4.64) yields

$$\begin{aligned} \frac{\partial b}{\partial \breve{\boldsymbol{A}}^*} = & -\boldsymbol{R}_{\boldsymbol{x},0}^{-1}\breve{\boldsymbol{A}}_0 - \frac{1}{t}\begin{bmatrix} \boldsymbol{I}_N & \mathbf{0}_{N\times L} \end{bmatrix} \boldsymbol{F}^{-1}(\breve{\boldsymbol{A}}, \tilde{\sigma}) \begin{bmatrix} \mathbf{0}_{N\times L} \\ \boldsymbol{I}_L \end{bmatrix} \\ & - \frac{1}{t}\sum_{l=1}^{L}\left(\frac{\frac{1}{2}\hat{\boldsymbol{a}}_l\boldsymbol{e}_l^T}{\Re\{\hat{\boldsymbol{a}}_l^H\breve{\boldsymbol{a}}_l\} - \kappa_l} - \frac{\nu_l^2\breve{\boldsymbol{a}}_l\boldsymbol{e}_l^T}{\kappa_l^2 - \nu_l^2\left\|\breve{\boldsymbol{a}}_l\right\|_2^2}\right) \end{aligned} \tag{A.69}$$

Generally, if

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{M}_{11} & \boldsymbol{M}_{12} \\ \boldsymbol{M}_{21} & \boldsymbol{M}_{22} \end{bmatrix} \tag{A.70}$$

is a square non-singular matrix and if $\boldsymbol{M}_{11}^{-1}$ and $\boldsymbol{M}_{22}^{-1}$ exist, then [vT02]

$$\boldsymbol{M}^{-1} = \begin{bmatrix} \boldsymbol{S}_{c1}^{-1} & -\boldsymbol{S}_{c1}^{-1}\boldsymbol{M}_{12}\boldsymbol{M}_{22}^{-1} \\ -\boldsymbol{S}_{c2}^{-1}\boldsymbol{M}_{21}\boldsymbol{M}_{11}^{-1} & \boldsymbol{S}_{c2}^{-1} \end{bmatrix}, \tag{A.71}$$

where

$$\begin{aligned}\boldsymbol{S}_{c1} &= \boldsymbol{M}_{11} - \boldsymbol{M}_{12}\boldsymbol{M}_{22}^{-1}\boldsymbol{M}_{21}\\ \boldsymbol{S}_{c2} &= \boldsymbol{M}_{22} - \boldsymbol{M}_{21}\boldsymbol{M}_{11}^{-1}\boldsymbol{M}_{12}.\end{aligned} \tag{A.72}$$

Consequently,

$$\begin{bmatrix} \boldsymbol{I}_N & \boldsymbol{0}_{N\times L} \end{bmatrix} \boldsymbol{F}^{-1}(\breve{\boldsymbol{A}}, \tilde{\sigma}) \begin{bmatrix} \boldsymbol{0}_{N\times L} \\ \boldsymbol{I}_L \end{bmatrix} = -\boldsymbol{\Delta}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}^{-1}(\breve{\boldsymbol{A}}, \tilde{\sigma})\breve{\boldsymbol{A}}, \tag{A.73}$$

where

$$\boldsymbol{\Delta}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}(\breve{\boldsymbol{A}}, \tilde{\sigma}) = \hat{\boldsymbol{R}}_{\boldsymbol{x}} - \breve{\boldsymbol{A}}\breve{\boldsymbol{A}}^H - \tilde{\sigma}\boldsymbol{I}_N. \tag{A.74}$$

Hence,

$$\frac{\partial b}{\partial \breve{\boldsymbol{A}}^*} = -\boldsymbol{R}_{\boldsymbol{x},0}^{-1}\breve{\boldsymbol{A}}_0 + \frac{1}{t}\boldsymbol{\Delta}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}^{-1}(\breve{\boldsymbol{A}}, \tilde{\sigma})\breve{\boldsymbol{A}} - \frac{1}{t}\sum_{l=1}^{L}\left(\frac{\frac{1}{2}\hat{\boldsymbol{a}}_l\boldsymbol{e}_l^T}{\Re\{\hat{\boldsymbol{a}}_l^H\breve{\boldsymbol{a}}_l\} - \kappa_l} - \frac{\nu_l^2\breve{\boldsymbol{a}}_l\boldsymbol{e}_l^T}{\kappa_l^2 - \nu_l^2\left\|\breve{\boldsymbol{a}}_l\right\|_2^2}\right). \tag{A.75}$$

From (4.64) and (A.71), we also obtain

$$\frac{\partial b}{\partial \tilde{\sigma}} = -\mathrm{tr}\left\{\boldsymbol{R}_{\boldsymbol{x},0}^{-1}\right\} + \frac{1}{t}\mathrm{tr}\left\{\boldsymbol{\Delta}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}^{-1}(\breve{\boldsymbol{A}}, \tilde{\sigma})\right\} - \frac{1}{t\tilde{\sigma}}. \tag{A.76}$$

Moreover, standard derivative rules give

$$\frac{\partial b}{\partial \kappa_l} = \frac{1}{t}\left(\frac{1}{\Re\{\hat{\boldsymbol{a}}_l^H\breve{\boldsymbol{a}}_l\} - \kappa_l} - \frac{2\kappa_l}{\kappa_l^2 - \nu_l^2\left\|\breve{\boldsymbol{a}}_l\right\|_2^2} - \frac{1}{\kappa_l}\right), \quad l = 1, \ldots, L. \tag{A.77}$$

It can be verified straightforwardly that (A.68) and (A.75)–(A.77) allow computing the gradient (A.66) with complexity $\mathcal{O}(N^3)$.

Let us now determine the Hessian (A.67). To compute the second order derivatives of the barrier function with respect to the real and imaginary parts of $\breve{\boldsymbol{A}}$, we use [Bra83]

$$\frac{\partial^2 b}{\partial \mathrm{vec}\{\Re\{\breve{\boldsymbol{A}}\}\}\partial \mathrm{vec}^T\{\Re\{\breve{\boldsymbol{A}}\}\}} = 2\Re\left\{\frac{\partial^2 b}{\partial \mathrm{vec}\{\breve{\boldsymbol{A}}^*\}\partial \mathrm{vec}^T\{\breve{\boldsymbol{A}}^*\}} + \frac{\partial^2 b}{\partial \mathrm{vec}\{\breve{\boldsymbol{A}}^*\}\partial \mathrm{vec}^T\{\breve{\boldsymbol{A}}\}}\right\} \tag{A.78}$$

$$\frac{\partial^2 b}{\partial \mathrm{vec}\{\Re\{\breve{\boldsymbol{A}}\}\}\partial \mathrm{vec}^T\{\Im\{\breve{\boldsymbol{A}}\}\}} = 2\Im\left\{\frac{\partial^2 b}{\partial \mathrm{vec}\{\breve{\boldsymbol{A}}^*\}\partial \mathrm{vec}^T\{\breve{\boldsymbol{A}}^*\}} - \frac{\partial^2 b}{\partial \mathrm{vec}\{\breve{\boldsymbol{A}}^*\}\partial \mathrm{vec}^T\{\breve{\boldsymbol{A}}\}}\right\} \tag{A.79}$$

$$\frac{\partial^2 b}{\partial \mathrm{vec}\{\Im\{\breve{\boldsymbol{A}}\}\}\partial \mathrm{vec}^T\{\Im\{\breve{\boldsymbol{A}}\}\}} = 2\Re\left\{\frac{\partial^2 b}{\partial \mathrm{vec}\{\breve{\boldsymbol{A}}^*\}\partial \mathrm{vec}^T\{\breve{\boldsymbol{A}}\}} - \frac{\partial^2 b}{\partial \mathrm{vec}\{\breve{\boldsymbol{A}}^*\}\partial \mathrm{vec}^T\{\breve{\boldsymbol{A}}^*\}}\right\}. \tag{A.80}$$

Moreover, we use that [vT02]

$$\frac{\partial \boldsymbol{M}^{-1}(\mu)}{\partial \mu} = -\boldsymbol{M}^{-1}(\mu)\frac{\partial \boldsymbol{M}(\mu)}{\partial \mu}\boldsymbol{M}^{-1}(\mu), \tag{A.81}$$

where $\boldsymbol{M}(\mu)$ is a matrix-valued function. For the second term on the right-hand side of (A.75), we obtain

$$\frac{\partial}{\partial\big[\breve{\boldsymbol{A}}^*\big]_{n,l}}\left\{\boldsymbol{\Delta}^{-1}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}(\breve{\boldsymbol{A}},\tilde{\sigma})\breve{\boldsymbol{A}}\right\} = \boldsymbol{\Delta}^{-1}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}(\breve{\boldsymbol{A}},\tilde{\sigma})\breve{\boldsymbol{A}}\boldsymbol{e}_l\boldsymbol{e}_n^T\boldsymbol{\Delta}^{-1}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}(\breve{\boldsymbol{A}},\tilde{\sigma})\breve{\boldsymbol{A}}. \tag{A.82}$$

Equation (5.28) gives

$$\begin{aligned}
&\mathrm{vec}\left\{\boldsymbol{\Delta}^{-1}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}(\breve{\boldsymbol{A}},\tilde{\sigma})\breve{\boldsymbol{A}}\boldsymbol{e}_l\boldsymbol{e}_n^T\boldsymbol{\Delta}^{-1}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}(\breve{\boldsymbol{A}},\tilde{\sigma})\breve{\boldsymbol{A}}\right\}\\
&\qquad = \left(\left[\boldsymbol{\Delta}^{-1}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}(\breve{\boldsymbol{A}},\tilde{\sigma})\breve{\boldsymbol{A}}\right]^T \otimes \left[\boldsymbol{\Delta}^{-1}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}(\breve{\boldsymbol{A}},\tilde{\sigma})\breve{\boldsymbol{A}}\right]\right)\mathrm{vec}\{\underset{(L)}{\boldsymbol{e}_l}\,\underset{(N)}{\boldsymbol{e}_n^T}\}\\
&\qquad = \left(\left[\boldsymbol{\Delta}^{-1}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}(\breve{\boldsymbol{A}},\tilde{\sigma})\breve{\boldsymbol{A}}\right]^T \otimes \left[\boldsymbol{\Delta}^{-1}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}(\breve{\boldsymbol{A}},\tilde{\sigma})\breve{\boldsymbol{A}}\right]\right)\boldsymbol{U}_{N,L}\mathrm{vec}\{\underset{(N)}{\boldsymbol{e}_n}\,\underset{(L)}{\boldsymbol{e}_l^T}\},
\end{aligned} \tag{A.83}$$

where [Bre78, Bre79]

$$\boldsymbol{U}_{N,L} = \sum_{n=1}^{N}\sum_{l=1}^{L}\big(\underset{(N)}{\boldsymbol{e}_n}\,\underset{(L)}{\boldsymbol{e}_l^T}\big)\otimes\big(\underset{(L)}{\boldsymbol{e}_l}\,\underset{(N)}{\boldsymbol{e}_n^T}\big) \tag{A.84}$$

is an $LN \times LN$ permutation matrix, and parenthetical underscores have been used to specify the vector-dimensions. Furthermore, standard derivative rules yield

$$\frac{\partial}{\partial\big[\breve{\boldsymbol{A}}^*\big]_{n,l}}\left\{\sum_{m=1}^{L}\frac{-\frac{1}{2}\hat{\boldsymbol{a}}_m\boldsymbol{e}_m^T}{\Re\{\hat{\boldsymbol{a}}_m^H\breve{\boldsymbol{a}}_m\}-\kappa_m}\right\} = \frac{\frac{1}{4}\hat{\boldsymbol{a}}_l\hat{\boldsymbol{a}}_l^T\boldsymbol{e}_n\boldsymbol{e}_l^T}{\big(\Re\{\hat{\boldsymbol{a}}_l^H\breve{\boldsymbol{a}}_l\}-\kappa_l\big)^2} \tag{A.85}$$

$$\frac{\partial}{\partial\big[\breve{\boldsymbol{A}}^*\big]_{n,l}}\left\{\sum_{m=1}^{L}\frac{\nu_m^2\breve{\boldsymbol{a}}_m\boldsymbol{e}_m^T}{\kappa_m^2-\nu_m^2\,\|\breve{\boldsymbol{a}}_m\|_2^2}\right\} = \frac{\nu_l^4\breve{\boldsymbol{a}}_l\breve{\boldsymbol{a}}_l^T\boldsymbol{e}_n\boldsymbol{e}_l^T}{\big(\kappa_l^2-\nu_l^2\,\|\breve{\boldsymbol{a}}_l\|_2^2\big)^2}. \tag{A.86}$$

Vectorizing the latter results gives

$$\mathrm{vec}\left\{\frac{\frac{1}{4}\hat{\boldsymbol{a}}_l\hat{\boldsymbol{a}}_l^T\boldsymbol{e}_n\boldsymbol{e}_l^T}{\big(\Re\{\hat{\boldsymbol{a}}_l^H\breve{\boldsymbol{a}}_l\}-\kappa_l\big)^2}\right\} = \frac{\frac{1}{4}\boldsymbol{e}_l\otimes\big(\hat{\boldsymbol{a}}_l\hat{\boldsymbol{a}}_l^T\big)}{\big(\Re\{\hat{\boldsymbol{a}}_l^H\breve{\boldsymbol{a}}_l\}-\kappa_l\big)^2}\boldsymbol{e}_n \tag{A.87}$$

$$\mathrm{vec}\left\{\frac{\nu_l^4\breve{\boldsymbol{a}}_l\breve{\boldsymbol{a}}_l^T\boldsymbol{e}_n\boldsymbol{e}_l^T}{\big(\kappa_l^2-\nu_l^2\,\|\breve{\boldsymbol{a}}_l\|_2^2\big)^2}\right\} = \frac{\nu_l^4\boldsymbol{e}_l\otimes\big(\breve{\boldsymbol{a}}_l\breve{\boldsymbol{a}}_l^T\big)}{\big(\kappa_l^2-\nu_l^2\,\|\breve{\boldsymbol{a}}_l\|_2^2\big)^2}\boldsymbol{e}_n. \tag{A.88}$$

Equations (A.75) and (A.82)–(A.88) yield

$$\begin{aligned}
\frac{\partial^2 b}{\partial\mathrm{vec}\{\breve{\boldsymbol{A}}^*\}\partial\mathrm{vec}^T\{\breve{\boldsymbol{A}}^*\}} &= \frac{1}{t}\left(\left[\boldsymbol{\Delta}^{-1}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}(\breve{\boldsymbol{A}},\tilde{\sigma})\breve{\boldsymbol{A}}\right]^T \otimes \left[\boldsymbol{\Delta}^{-1}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}(\breve{\boldsymbol{A}},\tilde{\sigma})\breve{\boldsymbol{A}}\right]\right)\boldsymbol{U}_{N\times L}\\
&\quad + \frac{1}{t}\left[\boldsymbol{\Omega}_1,\ldots,\boldsymbol{\Omega}_L\right],
\end{aligned} \tag{A.89}$$

where

$$\boldsymbol{\Omega}_l = \frac{\frac{1}{4}\boldsymbol{e}_l \otimes \left(\hat{\boldsymbol{a}}_l \hat{\boldsymbol{a}}_l^T\right)}{\left(\Re\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{a}}_l\} - \kappa_l\right)^2} + \frac{\nu_l^4 \boldsymbol{e}_l \otimes \left(\breve{\boldsymbol{a}}_l \breve{\boldsymbol{a}}_l^T\right)}{\left(\kappa_l^2 - \nu_l^2 \|\breve{\boldsymbol{a}}_l\|_2^2\right)^2}, \quad l = 1, \ldots, L. \tag{A.90}$$

Similarly, it can be shown that

$$\begin{aligned}\frac{\partial^2 b}{\partial \mathrm{vec}\{\breve{\boldsymbol{A}}^*\} \partial \mathrm{vec}^T\{\breve{\boldsymbol{A}}\}} &= \frac{1}{t}\left[\boldsymbol{I}_L + \breve{\boldsymbol{A}}^H \boldsymbol{\Delta}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}^{-1}(\breve{\boldsymbol{A}}, \tilde{\sigma})\breve{\boldsymbol{A}}\right]^T \otimes \boldsymbol{\Delta}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}^{-1}(\breve{\boldsymbol{A}}, \tilde{\sigma}) \\ &\quad + \frac{1}{t}\left[\tilde{\boldsymbol{\Omega}}_1, \ldots, \tilde{\boldsymbol{\Omega}}_L\right],\end{aligned} \tag{A.91}$$

where

$$\tilde{\boldsymbol{\Omega}}_l = \frac{\frac{1}{4}\boldsymbol{e}_l \otimes \left(\hat{\boldsymbol{a}}_l \hat{\boldsymbol{a}}_l^H\right)}{\left(\Re\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{a}}_l\} - \kappa_l\right)^2} + \frac{\nu_l^4 \boldsymbol{e}_l \otimes \left(\breve{\boldsymbol{a}}_l \breve{\boldsymbol{a}}_l^H\right)}{\left(\kappa_l^2 - \nu_l^2 \|\breve{\boldsymbol{a}}_l\|_2^2\right)^2} + \frac{\nu_l^2 \boldsymbol{e}_l \otimes \boldsymbol{I}_N}{\kappa_l^2 - \nu_l^2 \|\breve{\boldsymbol{a}}_l\|_2^2}, \quad l = 1, \ldots, L. \tag{A.92}$$

Equations (A.78)–(A.80), (A.89), and (A.91) determine the second-order derivatives of the barrier function with respect to the real and imaginary parts of $\breve{\boldsymbol{A}}$.

To compute the remaining blocks of the Hessian, we use [Bra83]

$$\frac{\partial^2 b}{\partial \tilde{\sigma} \partial \Re\{\breve{\boldsymbol{A}}\}} = 2\Re\left\{\frac{\partial^2 b}{\partial \tilde{\sigma} \partial \breve{\boldsymbol{A}}^*}\right\} \qquad \frac{\partial^2 b}{\partial \tilde{\sigma} \partial \Im\{\breve{\boldsymbol{A}}\}} = 2\Im\left\{\frac{\partial^2 b}{\partial \tilde{\sigma} \partial \breve{\boldsymbol{A}}^*}\right\} \tag{A.93}$$

$$\frac{\partial^2 b}{\partial \kappa_l \partial \Re\{\breve{\boldsymbol{A}}\}} = 2\Re\left\{\frac{\partial^2 b}{\partial \kappa_l \partial \breve{\boldsymbol{A}}^*}\right\} \qquad \frac{\partial^2 b}{\partial \kappa_l \partial \Im\{\breve{\boldsymbol{A}}\}} = 2\Im\left\{\frac{\partial^2 b}{\partial \kappa_l \partial \breve{\boldsymbol{A}}^*}\right\}. \tag{A.94}$$

Equations (A.75) and (A.81) give

$$\frac{\partial^2 b}{\partial \tilde{\sigma} \partial \breve{\boldsymbol{A}}^*} = \frac{1}{t}\boldsymbol{\Delta}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}^{-2}(\breve{\boldsymbol{A}}, \tilde{\sigma})\breve{\boldsymbol{A}}. \tag{A.95}$$

From (A.75), we also obtain

$$\frac{\partial^2 b}{\partial \kappa_l \partial \breve{\boldsymbol{A}}^*} = -\frac{1}{t}\left(\frac{\frac{1}{2}\hat{\boldsymbol{a}}_l \boldsymbol{e}_l^T}{\left(\Re\{\hat{\boldsymbol{a}}_l^H \breve{\boldsymbol{a}}_l\} - \kappa_l\right)^2} + \frac{2\kappa_l \nu_l^2 \breve{\boldsymbol{a}}_l \boldsymbol{e}_l^T}{\left(\kappa_l^2 - \nu_l^2 \|\breve{\boldsymbol{a}}_l\|_2^2\right)^2}\right), \quad l = 1, \ldots, L. \tag{A.96}$$

Moreover, (A.76) and (A.81) result in

$$\frac{\partial^2 b}{\partial \tilde{\sigma}^2} = \frac{1}{t}\left(\mathrm{tr}\left\{\boldsymbol{\Delta}_{\hat{\boldsymbol{R}}_{\boldsymbol{x}}}^{-2}(\breve{\boldsymbol{A}}, \tilde{\sigma})\right\} + \frac{1}{\tilde{\sigma}^2}\right). \tag{A.97}$$

It can be seen easily that

$$\frac{\partial^2 b}{\partial \tilde{\sigma} \partial \kappa_l} = 0, \quad l = 1, \ldots, L. \tag{A.98}$$

Finally, the lower-right block of the Hessian (A.67) follows from

$$\frac{\partial^2 b}{\partial \kappa_{l_1} \partial \kappa_{l_2}} = \frac{\delta_{l_1,l_2}}{t}\left(\frac{1}{\left(\Re\{\hat{\boldsymbol{a}}_{l_1}^H \breve{\boldsymbol{a}}_{l_1}\} - \kappa_{l_1}\right)^2} + \frac{2\left(\kappa_{l_1}^2 + \nu_{l_1}^2 \|\breve{\boldsymbol{a}}_{l_1}\|_2^2\right)}{\left(\kappa_{l_1}^2 - \nu_{l_1}^2 \|\breve{\boldsymbol{a}}_{l_1}\|_2^2\right)^2} + \frac{1}{\kappa_{l_1}^2}\right), \tag{A.99}$$

where δ_{l_1,l_2} is the Kronecker delta. Using the above results for the second-order derivatives, and taking into account its symmetry, the Hessian (A.67) can be computed with the complexity $\mathcal{O}(N^3 + L^2N^2)$.

A.7 Computational complexity of (5.76)

The barrier function corresponding to (5.76) can be expressed as [BV04]

$$b(\boldsymbol{W},\boldsymbol{\kappa},t) = \mathrm{vec}^H\{\boldsymbol{W}\}\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\mathrm{vec}\{\boldsymbol{W}\} - \frac{1}{t}\sum_{k=1}^{N_f}\Bigg(\log\left(\Re\left\{\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\right\} - \kappa_k - \gamma(f_k)\right)$$
$$+ \log\left(\kappa_k^2 - \epsilon_1^2(f_k+f_c)\left\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\right\|_2^2\right) + \log\kappa_k\Bigg), \tag{A.100}$$

where $\boldsymbol{\kappa} = [\kappa_1,\ldots,\kappa_{N_f}]^T$ is an $N_f \times 1$ vector of positive slack variables, and $t > 0$. Let

$$\boldsymbol{v} = \begin{bmatrix} \mathrm{vec}\{\Re\{\boldsymbol{W}\}\} \\ \mathrm{vec}\{\Im\{\boldsymbol{W}\}\} \\ \boldsymbol{\kappa} \end{bmatrix} \tag{A.101}$$

denote the $(2M_TN+N_f)\times 1$ vector of real-valued optimization variables. The gradient and Hessian of the barrier function can be written as

$$\frac{\partial b}{\partial \boldsymbol{v}} = \begin{bmatrix} \frac{\partial b}{\partial \mathrm{vec}\{\Re\{\boldsymbol{W}\}\}} \\ \frac{\partial b}{\partial \mathrm{vec}\{\Im\{\boldsymbol{W}\}\}} \\ \frac{\partial b}{\partial \boldsymbol{\kappa}} \end{bmatrix} \tag{A.102}$$

and

$$\frac{\partial^2 b}{\partial \boldsymbol{v}\partial \boldsymbol{v}^T} = \begin{bmatrix} \frac{\partial^2 b}{\partial \mathrm{vec}\{\Re\{\boldsymbol{W}\}\}\partial \mathrm{vec}^T\{\Re\{\boldsymbol{W}\}\}} & \frac{\partial^2 b}{\partial \mathrm{vec}\{\Re\{\boldsymbol{W}\}\}\partial \mathrm{vec}^T\{\Im\{\boldsymbol{W}\}\}} & \frac{\partial^2 b}{\partial \mathrm{vec}\{\Re\{\boldsymbol{W}\}\}\partial \boldsymbol{\kappa}^T} \\ \frac{\partial^2 b}{\partial \mathrm{vec}\{\Im\{\boldsymbol{W}\}\}\partial \mathrm{vec}^T\{\Re\{\boldsymbol{W}\}\}} & \frac{\partial^2 b}{\partial \mathrm{vec}\{\Im\{\boldsymbol{W}\}\}\partial \mathrm{vec}^T\{\Im\{\boldsymbol{W}\}\}} & \frac{\partial^2 b}{\partial \mathrm{vec}\{\Im\{\boldsymbol{W}\}\}\partial \boldsymbol{\kappa}^T} \\ \frac{\partial^2 b}{\partial \boldsymbol{\kappa}\partial \mathrm{vec}^T\{\Re\{\boldsymbol{W}\}\}} & \frac{\partial^2 b}{\partial \boldsymbol{\kappa}\partial \mathrm{vec}^T\{\Im\{\boldsymbol{W}\}\}} & \frac{\partial^2 b}{\partial \boldsymbol{\kappa}\partial \boldsymbol{\kappa}^T} \end{bmatrix}. \tag{A.103}$$

To obtain the derivatives with respect to the real and imaginary parts of $\boldsymbol{W}$, we use [Bra83]

$$\frac{\partial b}{\partial \mathrm{vec}\{\Re\{\boldsymbol{W}\}\}} = 2\Re\left\{\frac{\partial b}{\partial \mathrm{vec}\{\boldsymbol{W}^*\}}\right\} \tag{A.104}$$

$$\frac{\partial b}{\partial \mathrm{vec}\{\Im\{\boldsymbol{W}\}\}} = 2\Im\left\{\frac{\partial b}{\partial \mathrm{vec}\{\boldsymbol{W}^*\}}\right\} \tag{A.105}$$

$$\frac{\partial^2 b}{\partial \kappa_k \partial \mathrm{vec}\{\Re\{\boldsymbol{W}\}\}} = 2\Re\left\{\frac{\partial^2 b}{\partial \kappa_k \partial \mathrm{vec}\{\boldsymbol{W}^*\}}\right\} \tag{A.106}$$

$$\frac{\partial^2 b}{\partial \kappa_k \partial \mathrm{vec}\{\Im\{\boldsymbol{W}\}\}} = 2\Im\left\{\frac{\partial^2 b}{\partial \kappa_k \partial \mathrm{vec}\{\boldsymbol{W}^*\}}\right\} \tag{A.107}$$

$$\frac{\partial^2 b}{\partial \text{vec}\{\Re\{\boldsymbol{W}\}\}\partial \text{vec}^T\{\Re\{\boldsymbol{W}\}\}} = 2\Re\left\{\frac{\partial}{\partial \text{vec}\{\boldsymbol{W}^*\}}\left(\frac{\partial b}{\partial \text{vec}^T\{\boldsymbol{W}^*\}} + \frac{\partial b}{\partial \text{vec}^T\{\boldsymbol{W}\}}\right)\right\} \tag{A.108}$$

$$\frac{\partial^2 b}{\partial \text{vec}\{\Re\{\boldsymbol{W}\}\}\partial \text{vec}^T\{\Im\{\boldsymbol{W}\}\}} = 2\Im\left\{\frac{\partial}{\partial \text{vec}\{\boldsymbol{W}^*\}}\left(\frac{\partial b}{\partial \text{vec}^T\{\boldsymbol{W}^*\}} - \frac{\partial b}{\partial \text{vec}^T\{\boldsymbol{W}\}}\right)\right\} \tag{A.109}$$

$$\frac{\partial^2 b}{\partial \text{vec}\{\Im\{\boldsymbol{W}\}\}\partial \text{vec}^T\{\Im\{\boldsymbol{W}\}\}} = 2\Re\left\{\frac{\partial}{\partial \text{vec}\{\boldsymbol{W}^*\}}\left(\frac{\partial b}{\partial \text{vec}^T\{\boldsymbol{W}\}} - \frac{\partial b}{\partial \text{vec}^T\{\boldsymbol{W}^*\}}\right)\right\}. \tag{A.110}$$

Using (5.28), we obtain

$$\frac{\partial b}{\partial \text{vec}\{\boldsymbol{W}^*\}} = \underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\text{vec}\{\boldsymbol{W}\} - \frac{1}{t}\sum_{k=1}^{N_f}\boldsymbol{d}_{\text{cs}}(f_k)\otimes\left(\frac{\boldsymbol{1}_N}{2\nu_1(f_k)} - \frac{\epsilon_1^2(f_k+f_c)\boldsymbol{W}\boldsymbol{d}_{\text{cs}}^*(f_k)}{\nu_2(f_k)}\right), \tag{A.111}$$

where

$$\begin{aligned}\nu_1(f_k) &= \Re\left\{\boldsymbol{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\text{cs}}(f_k)\right\} - \kappa_k - \gamma(f_k), \quad k=1,\ldots,N_f\\ \nu_2(f_k) &= \kappa_k^2 - \epsilon_1^2(f_k+f_c)\left\|\boldsymbol{W}^*\boldsymbol{d}_{\text{cs}}(f_k)\right\|_2^2, \quad k=1,\ldots,N_f.\end{aligned} \tag{A.112}$$

Moreover, standard derivative rules yield

$$\frac{\partial b}{\partial \kappa_k} = \frac{1}{t}\left(\frac{1}{\nu_1(f_k)} - \frac{2\kappa_k}{\nu_2(f_k)} - \frac{1}{\kappa_k}\right), \quad k=1,\ldots,N_f. \tag{A.113}$$

Equations (A.111) and (A.113) allow computing the gradient (A.102) with complexity $\mathcal{O}(M_T^2N^2 + M_TNN_f)$.

Using (5.28) and (5.111), we obtain from (A.111)

$$\begin{aligned}\frac{\partial^2 b}{\partial \text{vec}\{\boldsymbol{W}^*\}\partial \text{vec}^T\{\boldsymbol{W}^*\}} = \frac{1}{t}\sum_{k=1}^{N_f}\left(\boldsymbol{d}_{\text{cs}}(f_k)\boldsymbol{d}_{\text{cs}}^T(f_k)\right)\otimes\Bigg(&\frac{\boldsymbol{1}_N\boldsymbol{1}_N^T}{4\nu_1^2(f_k)}\\ &+\frac{\epsilon_1^4(f_k+f_c)\boldsymbol{W}\boldsymbol{d}_{\text{cs}}^*(f_k)\boldsymbol{d}_{\text{cs}}^H(f_k)\boldsymbol{W}^T}{\nu_2^2(f_k)}\Bigg).\end{aligned} \tag{A.114}$$

Similarly, it can be shown that

$$\begin{aligned}\frac{\partial^2 b}{\partial \text{vec}\{\boldsymbol{W}^*\}\partial \text{vec}^T\{\boldsymbol{W}\}} = \underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}} + \frac{1}{t}\sum_{k=1}^{N_f}\left(\boldsymbol{d}_{\text{cs}}(f_k)\boldsymbol{d}_{\text{cs}}^H(f_k)\right)\otimes\Bigg(&\frac{\boldsymbol{1}_N\boldsymbol{1}_N^T}{4\nu_1^2(f_k)}\\ +\frac{\epsilon_1^4(f_k+f_c)\boldsymbol{W}\boldsymbol{d}_{\text{cs}}^*(f_k)\boldsymbol{d}_{\text{cs}}^T(f_k)\boldsymbol{W}^H}{\nu_2^2(f_k)} &+ \frac{\epsilon_1^2(f_k+f_c)\boldsymbol{I}_N}{\nu_2(f_k)}\Bigg).\end{aligned} \tag{A.115}$$

Equations (A.108)–(A.110), (A.114), and (A.115) determine the second-order derivatives of the barrier function with respect to the real and imaginary parts of $\boldsymbol{W}$.

The remaining blocks of the Hessian can be computed as follows. Equation (A.111) gives

$$\frac{\partial^2 b}{\partial \kappa_k \partial \mathrm{vec}\{\boldsymbol{W}^*\}} = -\frac{1}{t}\boldsymbol{d}_{\mathrm{cs}}(f_k) \otimes \left(\frac{\mathbf{1}_N}{2\nu_1^2(f_k)} + \frac{2\kappa_k \epsilon_1^2(f_k+f_c)\boldsymbol{W}\boldsymbol{d}^*_{\mathrm{cs}}(f_k)}{\nu_2^2(f_k)}\right). \tag{A.116}$$

Finally, the lower-right block of the Hessian (A.103) is determined by

$$\frac{\partial^2 b}{\partial \kappa_{k_1} \partial \kappa_{k_2}} = \frac{\delta_{k_1,k_2}}{t}\left(\frac{1}{\nu_1^2(f_{k_1})} - \frac{2}{\nu_2(f_{k_1})} + \frac{4\kappa_{k_1}^2}{\nu_2^2(f_{k_1})} + \frac{1}{\kappa_{k_1}^2}\right), \tag{A.117}$$

where δ_{k_1,k_2} is the Kronecker delta. The above results specify the Hessian of the barrier function uniquely. Exploiting that in (A.114)

$$\boldsymbol{d}_{\mathrm{cs}}(f_k)\boldsymbol{d}^T_{\mathrm{cs}}(f_k), \quad k = 1, \ldots, N_f \tag{A.118}$$

are Hankel matrices, and that in (A.115)

$$\boldsymbol{d}_{\mathrm{cs}}(f_k)\boldsymbol{d}^H_{\mathrm{cs}}(f_k), \quad k = 1, \ldots, N_f \tag{A.119}$$

are Toeplitz matrices, the Hessian (A.103) can be computed with the complexity $\mathcal{O}(M_T^2 N^2 + M_T N^2 N_f)$.

A.8 Computational complexity of (5.90)

The barrier function corresponding to (5.90) can be expressed as [BV04][1]

$$\begin{aligned} b(\boldsymbol{W}, \boldsymbol{\kappa}, t) = {} & \mathrm{vec}^H\{\boldsymbol{W}\}\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\mathrm{vec}\{\boldsymbol{W}\} \\ & - \frac{1}{t}\sum_{k=1}^{N_f}\Bigg(\log\left((\zeta(f_k) - \kappa_k)^2 - \epsilon_1^2(f_k+f_c)\left\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\right\|_2^2\right) \\ & + \log\left(\zeta(f_k) - \kappa_k\right) + \log\left(\kappa_k^2 - \left|\mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k) - 1\right|^2\right) + \log \kappa_k\Bigg), \end{aligned} \tag{A.120}$$

where $\boldsymbol{\kappa} = [\kappa_1, \ldots, \kappa_{N_f}]^T$ is an $N_f \times 1$ vector of positive slack variables, and $t > 0$. Defining the vector of optimization variables as in (A.101), the gradient and the Hessian of the barrier function can be expressed as in (A.102) and (A.103), respectively. To obtain the derivatives with respect to the real and imaginary parts of $\boldsymbol{W}$, we use (A.104)–(A.110). Using (5.28), we obtain

$$\begin{aligned} \frac{\partial b}{\partial \mathrm{vec}\{\boldsymbol{W}^*\}} = {} & \underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\mathrm{vec}\{\boldsymbol{W}\} \\ & + \frac{1}{t}\sum_{k=1}^{N_f}\boldsymbol{d}_{\mathrm{cs}}(f_k) \otimes \left(\frac{\epsilon_1^2(f_k+f_c)\boldsymbol{W}\boldsymbol{d}^*_{\mathrm{cs}}(f_k)}{\nu_1(f_k)} + \frac{\nu_2^*(f_k)\mathbf{1}_N}{\kappa_k^2 - |\nu_2(f_k)|^2}\right), \end{aligned} \tag{A.121}$$

---

[1] In this appendix, the notation of Appendix A.7 is partially re-defined for notational simplicity.

where we defined

$$\begin{aligned}\nu_1(f_k) &= \left(\zeta(f_k)-\kappa_k\right)^2-\epsilon_1^2(f_k+f_c)\left\|\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)\right\|_2^2, \quad k=1,\dots,N_f\\ \nu_2(f_k) &= \mathbf{1}_N^T\boldsymbol{W}^*\boldsymbol{d}_{\mathrm{cs}}(f_k)-1, \quad k=1,\dots,N_f.\end{aligned} \tag{A.122}$$

Standard derivative rules yield

$$\frac{\partial b}{\partial \kappa_k}=\frac{1}{t}\left(\frac{2(\zeta(f_k)-\kappa_k)}{\nu_1(f_k)}+\frac{1}{\zeta(f_k)-\kappa_k}-\frac{2\kappa_k}{\kappa_k^2-|\nu_2(f_k)|^2}-\frac{1}{\kappa_k}\right), \quad k=1,\dots,N_f. \tag{A.123}$$

Equations (A.111) and (A.113) allow computing the gradient of the barrier function with complexity $\mathcal{O}(M_T^2N^2+M_TNN_f)$.

Using (5.28) and (5.111), we obtain from (A.102)

$$\begin{aligned}\frac{\partial^2 b}{\partial \mathrm{vec}\{\boldsymbol{W}^*\}\partial \mathrm{vec}^T\{\boldsymbol{W}^*\}} &= \frac{1}{t}\sum_{k=1}^{N_f}\left(\boldsymbol{d}_{\mathrm{cs}}(f_k)\boldsymbol{d}_{\mathrm{cs}}^T(f_k)\right)\otimes\\ &\left(\frac{\epsilon_1^4(f_k+f_c)\boldsymbol{W}\boldsymbol{d}_{\mathrm{cs}}^*(f_k)\boldsymbol{d}_{\mathrm{cs}}^H(f_k)\boldsymbol{W}^T}{\nu_1^2(f_k)}+\frac{\left(\nu_2^2(f_k)\right)^*\mathbf{1}_N\mathbf{1}_N^T}{\left(\kappa_k^2-|\nu_2(f_k)|^2\right)^2}\right).\end{aligned} \tag{A.124}$$

Similarly, it can be shown that

$$\begin{aligned}\frac{\partial^2 b}{\partial \mathrm{vec}\{\boldsymbol{W}^*\}\partial \mathrm{vec}^T\{\boldsymbol{W}\}} &= \underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}+\frac{1}{t}\sum_{k=1}^{N_f}\left(\boldsymbol{d}_{\mathrm{cs}}(f_k)\boldsymbol{d}_{\mathrm{cs}}^H(f_k)\right)\otimes\\ &\left(\frac{\epsilon_1^4(f_k+f_c)\boldsymbol{W}\boldsymbol{d}_{\mathrm{cs}}^*(f_k)\boldsymbol{d}_{\mathrm{cs}}^T(f_k)\boldsymbol{W}^H}{\nu_1^2(f_k)}+\frac{\epsilon_1^2(f_k+f_c)\boldsymbol{I}_N}{\nu_1(f_k)}\right.\\ &\left.+\frac{|\nu_2(f_k)|^2\mathbf{1}_N\mathbf{1}_N^T}{\left(\kappa_k^2-|\nu_2(f_k)|^2\right)^2}\right).\end{aligned} \tag{A.125}$$

Equations (A.124) and (A.125) determine the second-order derivatives of the barrier function with respect to the real and imaginary parts of $\boldsymbol{W}$.

The remaining blocks of the Hessian can be computed as follows. Equation (A.121) gives

$$\frac{\partial^2 b}{\partial \kappa_k \partial \mathrm{vec}\{\boldsymbol{W}^*\}}=\frac{\boldsymbol{d}_{\mathrm{cs}}(f_k)}{t}\otimes\left(\frac{2\epsilon_1^2(f_k+f_c)(\zeta(f_k)-\kappa_k)\boldsymbol{W}\boldsymbol{d}_{\mathrm{cs}}^*(f_k)}{\nu_1^2(f_k)}-\frac{2\kappa_k\nu_2^*(f_k)\mathbf{1}_N}{\left(\kappa_k^2-|\nu_2(f_k)|^2\right)^2}\right). \tag{A.126}$$

Furthermore, (A.123) gives

$$\begin{aligned}\frac{\partial^2 b}{\partial \kappa_{k_1}\partial \kappa_{k_2}} = \frac{\delta_{k_1,k_2}}{t}&\left(\frac{4\left(\zeta(f_{k_1})-\kappa_{k_1}\right)^2}{\nu_1^2(f_{k_1})}-\frac{2}{\nu_1(f_{k_1})}+\frac{4\kappa_{k_1}^2}{\left(\kappa_{k_1}^2-|\nu_2(f_{k_1})|^2\right)^2}-\frac{2}{\kappa_{k_1}^2-|\nu_2(f_{k_1})|^2}\right.\\ &\left.+\frac{1}{\left(\zeta(f_{k_1})-\kappa_{k_1}\right)^2}+\frac{1}{\kappa_{k_1}^2}\right).\end{aligned} \tag{A.127}$$

Using the above results for the second-order derivatives, and taking into account its symmetry, the Hessian of the barrier function can be computed with the complexity $\mathcal{O}(M_T^2N^2 + M_TN^2N_f)$.

A.9 Proof of Proposition 5.1

The problem in (5.90) is feasible if and only if all N_f constraints can be satisfied jointly. Let us first consider the constraint for a single frequency f_k. This constraint is feasible if and only if

$$\zeta(f_k) \geq \min_{\boldsymbol{W}} \left|\mathbf{1}_N^T \boldsymbol{W}^* \boldsymbol{d}_{\text{cs}}(f_k) - 1\right| + \epsilon_1(f_k)\|\boldsymbol{W}^*\boldsymbol{d}_{\text{cs}}(f_k)\|_2. \tag{A.128}$$

Since $\boldsymbol{W}^*\boldsymbol{d}_{\text{cs}}(f_k)$ is an arbitrary complex-valued vector, substituting $\boldsymbol{m}_k = \boldsymbol{W}^*\boldsymbol{d}_{\text{cs}}(f_k)$ in (A.128) yields

$$\zeta(f_k) \geq \min_{\boldsymbol{m}_k \in \mathbb{C}^{N\times 1}} \left|\mathbf{1}_N^T \boldsymbol{m}_k - 1\right| + \epsilon_1(f_k)\|\boldsymbol{m}_k\|_2. \tag{A.129}$$

The condition in (A.129) can be written equivalently as

$$\begin{aligned} \zeta(f_k) \geq \quad \min_{\mu_k, \boldsymbol{m}_k} \quad & |\mu_k - 1| + \epsilon_1(f_k)\|\boldsymbol{m}_k\|_2 \\ \text{s.t.} \quad & \mu_k = \mathbf{1}_N^T \boldsymbol{m}_k. \end{aligned} \tag{A.130}$$

For a fixed value of μ_k, the minimization with respect to $\boldsymbol{m}_k$ is equivalent to

$$\min_{\boldsymbol{m}_k} \ \boldsymbol{m}_k^H \boldsymbol{m}_k \quad \text{s.t.} \quad \mu_k = \mathbf{1}_N^T \boldsymbol{m}_k. \tag{A.131}$$

The optimum vector $\boldsymbol{m}_{k,\text{opt}}$ of (A.131) is given by

$$\boldsymbol{m}_{k,\text{opt}} = \mu_k \frac{\mathbf{1}_N}{N}. \tag{A.132}$$

Substituting (A.132) in (A.130) yields

$$\zeta(f_k) \geq \min_{\mu_k} \ |\mu_k - 1| + |\mu_k| \frac{\epsilon_1(f_k)}{\sqrt{N}}. \tag{A.133}$$

Let $\mu_{k,\text{opt}}$ denote the optimum point of (A.133). Clearly, $\mu_{k,\text{opt}}$ has to be real-valued. Furthermore, it can be verified easily that $\mu_{k,\text{opt}}$ has to be non-negative. The objective function in (A.133) is piecewise linear for real-valued non-negative μ_k. In particular, this function is linear on the intervals $0 \leq \mu_k \leq 1$ and $1 \leq \mu_k \leq \infty$. Checking the values of the objective function for $\mu_k = 0, 1,$ and ∞ leads to the solution

$$\mu_{k,\text{opt}} = 1, \tag{A.134}$$

where we have used that $\epsilon_1(f_k) < \sqrt{N}$. Hence,

$$\boldsymbol{m}_{k,\text{opt}} = \frac{\mathbf{1}_N}{N} \tag{A.135}$$

and the minimum value of (A.128) is given by $\epsilon_1(f_k)/\sqrt{N}$. The quiescent weight matrix (5.92) satisfies

$$\boldsymbol{m}_{k,\text{opt}} = \boldsymbol{W}_q^* \boldsymbol{d}_{\text{cs}}(f_k) \quad \forall\, k = 1, \ldots, N_f. \tag{A.136}$$

Hence,

$$\zeta(f_k) \geq \epsilon_1(f_k)/\sqrt{N} \tag{A.137}$$

is a tight lower bound for the feasibility of (5.90).

A.10 Computational complexity of (5.144)

The barrier function corresponding to (5.144) can be expressed as [BV04]

$$b(\boldsymbol{W}, \{\boldsymbol{Y}_m\}, \{\boldsymbol{Z}_m\}, t) = \text{vec}^H\{\boldsymbol{W}\} \underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}} \text{vec}\{\boldsymbol{W}\} - \frac{1}{t} \sum_{m=1}^{4N} (\log\det \boldsymbol{Y}_m + \log\det \boldsymbol{Z}_m)\,, \tag{A.138}$$

where $t > 0$. The barrier method involves solving a sequence of optimization problems

$$\begin{aligned} \min_{\boldsymbol{W}, \{\boldsymbol{Y}_m\}, \{\boldsymbol{Z}_m\}} \quad & b(\boldsymbol{W}, \{\boldsymbol{Y}_m\}, \{\boldsymbol{Z}_m\}, t) \\ \text{s.t.} \quad & \boldsymbol{f}_1(\boldsymbol{Y}_{(k_1+(k_2+3)/2)N+n}) - \boldsymbol{f}_2(\boldsymbol{Z}_{(k_1+(k_2+3)/2)N+n}, B_s, T_s) = \\ & \underline{\boldsymbol{I}}^T \left(\boldsymbol{W}^T + \boldsymbol{J}_{M_T} \boldsymbol{W}^H\right) \mathbf{1}_N - 2\underline{\boldsymbol{I}}^T \boldsymbol{q}_{\breve{\gamma}} - \\ & \tfrac{1}{2}\boldsymbol{f}_1\left((k_1 + jk_2)\boldsymbol{q}_{\breve{\epsilon}} \boldsymbol{e}_n^T \boldsymbol{W}^* + (k_1 - jk_2)\boldsymbol{W}^T \boldsymbol{e}_n \boldsymbol{q}_{\breve{\epsilon}}^H\right), \\ & k_1, k_2 \in \{-1, 1\}, \quad n = 1, \ldots, N \\ & \boldsymbol{Y}_m \in \mathcal{H}_+^{(M_T)},\ \boldsymbol{Z}_m \in \mathcal{H}_+^{(M_T-1)}, \quad m = 1, \ldots, 4N \end{aligned} \tag{A.139}$$

for different values of t.

Stacking the independent real and imaginary components of the $M_T \times M_T$ Hermitian matrix $\boldsymbol{Y}_m$ on top of each other (the order is irrelevant for our subsequent discussion), leads to a $M_T^2 \times 1$ real-valued vector $\boldsymbol{y}_m$. Similarly, $\boldsymbol{Z}_m$ can be represented by a real-valued $(M_T - 1)^2 \times 1$ vector $\boldsymbol{z}_m$. Then, the vector of optimization variables corresponding to (A.139) can be written as

$$\boldsymbol{v} = \begin{bmatrix} \mathrm{vec}\{\Re\{\boldsymbol{W}\}\} \\ \mathrm{vec}\{\Im\{\boldsymbol{W}\}\} \\ \boldsymbol{y} \\ \boldsymbol{z} \end{bmatrix} \in \mathbb{R}^{(2M_T N + 4M_T^2 N + 4(M_T-1)^2 N)\times 1}, \tag{A.140}$$

where

$$\begin{aligned} \boldsymbol{y} &= \left[\boldsymbol{y}_1^T, \dots, \boldsymbol{y}_{4N}^T\right]^T \in \mathbb{R}^{4M_T^2 N \times 1} \\ \boldsymbol{z} &= \left[\boldsymbol{z}_1^T, \dots, \boldsymbol{z}_{4N}^T\right]^T \in \mathbb{R}^{4(M_T-1)^2 N \times 1}. \end{aligned} \tag{A.141}$$

To solve (A.139) for a fixed value of t, Newton's method involves solving a sequence of quadratic approximations of this problem. The quadratic Taylor approximation of the barrier function about $\boldsymbol{v}$ is given by

$$b(\boldsymbol{v} + \Delta\boldsymbol{v}, t) \approx b(\boldsymbol{v}, t) + \boldsymbol{g}^T \Delta\boldsymbol{v} + \frac{1}{2}\Delta\boldsymbol{v}^T \boldsymbol{H} \Delta\boldsymbol{v}, \tag{A.142}$$

where $\boldsymbol{g}$ and $\boldsymbol{H}$ are the gradient and the Hessian of the barrier function, respectively. The gradient can be written as

$$\frac{\partial b}{\partial \boldsymbol{v}} = \begin{bmatrix} \frac{\partial b}{\partial \mathrm{vec}\{\Re\{\boldsymbol{W}\}\}} \\ \frac{\partial b}{\partial \mathrm{vec}\{\Im\{\boldsymbol{W}\}\}} \\ \frac{\partial b}{\partial \boldsymbol{y}} \\ \frac{\partial b}{\partial \boldsymbol{z}} \end{bmatrix}. \tag{A.143}$$

The Hessian of the barrier function is a block-diagonal matrix

$$\boldsymbol{H} = \frac{\partial^2 b}{\partial \boldsymbol{v} \partial \boldsymbol{v}^T} = \mathrm{diag}\left\{\boldsymbol{H}_{\boldsymbol{W}}, \boldsymbol{H}_{\boldsymbol{y}}, \boldsymbol{H}_{\boldsymbol{z}}\right\}, \tag{A.144}$$

where the first block is

$$\boldsymbol{H}_{\boldsymbol{W}} = \begin{bmatrix} \frac{\partial^2 b}{\partial \mathrm{vec}\{\Re\{\boldsymbol{W}\}\} \partial \mathrm{vec}^T\{\Re\{\boldsymbol{W}\}\}} & \frac{\partial^2 b}{\partial \mathrm{vec}\{\Re\{\boldsymbol{W}\}\} \partial \mathrm{vec}^T\{\Im\{\boldsymbol{W}\}\}} \\ \frac{\partial^2 b}{\partial \mathrm{vec}\{\Im\{\boldsymbol{W}\}\} \partial \mathrm{vec}^T\{\Re\{\boldsymbol{W}\}\}} & \frac{\partial^2 b}{\partial \mathrm{vec}\{\Im\{\boldsymbol{W}\}\} \partial \mathrm{vec}^T\{\Im\{\boldsymbol{W}\}\}} \end{bmatrix} \in \mathbb{R}^{2M_T N \times 2M_T N}, \tag{A.145}$$

and

$$\begin{aligned} \boldsymbol{H}_{\boldsymbol{y}} &= \mathrm{diag}\left\{\boldsymbol{H}_{\boldsymbol{y}_1}, \dots, \boldsymbol{H}_{\boldsymbol{y}_{4N}}\right\} \in \mathbb{R}^{4M_T^2 N \times 4M_T^2 N} \\ \boldsymbol{H}_{\boldsymbol{z}} &= \mathrm{diag}\left\{\boldsymbol{H}_{\boldsymbol{z}_1}, \dots, \boldsymbol{H}_{\boldsymbol{z}_{4N}}\right\} \in \mathbb{R}^{4(M_T-1)^2 N \times 4(M_T-1)^2 N} \end{aligned} \tag{A.146}$$

are itself block-diagonal matrices with

$$\begin{aligned} \boldsymbol{H}_{\boldsymbol{y}_m} &= \frac{\partial^2 b}{\partial \boldsymbol{y}_m \partial \boldsymbol{y}_m^T} \in \mathbb{R}^{M_T^2 \times M_T^2} \\ \boldsymbol{H}_{\boldsymbol{z}_m} &= \frac{\partial^2 b}{\partial \boldsymbol{z}_m \partial \boldsymbol{z}_m^T} \in \mathbb{R}^{(M_T-1)^2 \times (M_T-1)^2}. \end{aligned} \tag{A.147}$$

The constraints in (A.139) for $m = (k_1 + (k_2+3)/2)N + n$ can be written as

$$\boldsymbol{G}_m \boldsymbol{v} = \boldsymbol{b}, \tag{A.148}$$

where

$$\boldsymbol{G}_m = \left[\boldsymbol{G}_{\boldsymbol{W},m}, \underset{(4N)}{\boldsymbol{e}_m^T} \otimes \boldsymbol{G}_{\boldsymbol{y}}, \underset{(4N)}{\boldsymbol{e}_m^T} \otimes \boldsymbol{G}_{\boldsymbol{z}} \right] \in \mathbb{R}^{2M_T \times (2M_T N + 4M_T^2 N + 4(M_T-1)^2 N)}, \tag{A.149}$$

$\boldsymbol{G}_{\boldsymbol{y}}$ and $\boldsymbol{G}_{\boldsymbol{z}}$ are matrices of dimension $2M_T \times M_T^2$ and $2M_T \times (M_T-1)^2$, respectively,

$$\boldsymbol{b} = 2 \begin{bmatrix} \underline{\boldsymbol{I}}^T \Re\left\{\boldsymbol{q}_{\breve{\gamma}}\right\} \\ \underline{\boldsymbol{I}}^T \Im\left\{\boldsymbol{q}_{\breve{\gamma}}\right\} \end{bmatrix} \in \mathbb{R}^{2M_T \times 1}, \tag{A.150}$$

and parenthetical underscores have been used to specify the vector-dimensions. Hence, the constraints in (A.139) can be expressed as

$$\boldsymbol{G}\boldsymbol{v} = \mathbf{1}_{4N} \otimes \boldsymbol{b}, \tag{A.151}$$

where

$$\boldsymbol{G} = \left[\boldsymbol{G}_{\boldsymbol{W}}, \boldsymbol{I}_{4N} \otimes \boldsymbol{G}_{\boldsymbol{y}}, \boldsymbol{I}_{4N} \otimes \boldsymbol{G}_{\boldsymbol{z}}\right] \in \mathbb{R}^{8M_T N \times (2M_T N + 4M_T^2 N + 4(M_T-1)^2 N)}, \tag{A.152}$$

and

$$\boldsymbol{G}_{\boldsymbol{W}} = \begin{bmatrix} \boldsymbol{G}_{\boldsymbol{W},1} \\ \vdots \\ \boldsymbol{G}_{\boldsymbol{W},4N} \end{bmatrix} \in \mathbb{R}^{8M_T N \times 2M_T N}. \tag{A.153}$$

Allowing for infeasible starting points, the quadratic approximations of (A.139) used in Newton's method can be formulated as

$$\begin{aligned} \min_{\Delta \boldsymbol{v}} \quad & b(\boldsymbol{v}, t) + \boldsymbol{g}^T \Delta \boldsymbol{v} + \tfrac{1}{2} \Delta \boldsymbol{v}^T \boldsymbol{H} \Delta \boldsymbol{v} \\ \text{s.t.} \quad & \boldsymbol{G}(\boldsymbol{v} + \Delta \boldsymbol{v}) = \mathbf{1}_{4N} \otimes \boldsymbol{b}. \end{aligned} \tag{A.154}$$

The KKT system corresponding to (A.154) can be written as

$$\begin{bmatrix} \boldsymbol{H} & \boldsymbol{G}^T \\ \boldsymbol{G} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{v} \\ \boldsymbol{u} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{g} \\ \mathbf{1}_{4N} \otimes \boldsymbol{b} - \boldsymbol{G}\boldsymbol{v} \end{bmatrix}, \tag{A.155}$$

where $\boldsymbol{u}$ is an $8M_TN \times 1$ vector of Lagrange parameters. If $\left[\Delta\boldsymbol{v}^T, \boldsymbol{u}^T\right]^T$ solves (A.155), then $\boldsymbol{\Delta v}$ is an optimum point of (A.154) [BV04].

Let us first compute the gradient (A.143). Equations (A.104) and (A.105) give

$$\begin{aligned} \frac{\partial b}{\partial \mathrm{vec}\{\Re\{\boldsymbol{W}\}\}} &= 2\Re\left\{\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}} \mathrm{vec}\left\{\boldsymbol{W}\right\}\right\} \\ \frac{\partial b}{\partial \mathrm{vec}\{\Im\{\boldsymbol{W}\}\}} &= 2\Im\left\{\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}} \mathrm{vec}\left\{\boldsymbol{W}\right\}\right\}. \end{aligned} \tag{A.156}$$

Using (4.64), we obtain

$$\begin{aligned} \frac{\partial b}{\partial \left[\Re\{\boldsymbol{Y}_m\}\right]_{k_1,k_2}} &= -\frac{1}{t}\left(2-\delta_{k_1,k_2}\right)\Re\left\{\boldsymbol{e}_{k_1}^T \boldsymbol{Y}_m^{-1} \boldsymbol{e}_{k_2}\right\} \\ \frac{\partial b}{\partial \left[\Im\{\boldsymbol{Y}_m\}\right]_{k_1,k_2}} &= -\frac{2}{t}\Im\left\{\boldsymbol{e}_{k_1}^T \boldsymbol{Y}_m^{-1} \boldsymbol{e}_{k_2}\right\}. \end{aligned} \tag{A.157}$$

The latter results allow computing the derivative of the barrier function with respect to $\boldsymbol{y}_m$ with complexity $\mathcal{O}(M_T^3)$. The complexity to compute the gradient (A.143) is therefore $\mathcal{O}(M_T^2N^2 + M_T^3N)$.

Let us now determine the Hessian (A.144). Equations (A.108)–(A.110) give

$$\begin{aligned} \frac{\partial^2 b}{\partial \mathrm{vec}\left\{\Re\{\boldsymbol{W}\}\right\} \partial \mathrm{vec}^T\left\{\Re\{\boldsymbol{W}\}\right\}} &= 2\Re\left\{\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\right\} \\ \frac{\partial^2 b}{\partial \mathrm{vec}\left\{\Re\{\boldsymbol{W}\}\right\} \partial \mathrm{vec}^T\left\{\Im\{\boldsymbol{W}\}\right\}} &= -2\Im\left\{\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\right\} \\ \frac{\partial^2 b}{\partial \mathrm{vec}\left\{\Im\{\boldsymbol{W}\}\right\} \partial \mathrm{vec}^T\left\{\Im\{\boldsymbol{W}\}\right\}} &= 2\Re\left\{\underline{\hat{\boldsymbol{R}}}_{\boldsymbol{x}}\right\}. \end{aligned} \tag{A.158}$$

Moreover, we obtain from (A.157) and (A.81)

$$\begin{aligned} &\frac{\partial^2 b}{\partial \left[\Re\{\boldsymbol{Y}_m\}\right]_{k_1,k_2} \partial \left[\Re\{\boldsymbol{Y}_m\}\right]_{n_1,n_2}} \\ &\quad = \frac{2-\delta_{k_1,k_2}}{t\left(1+\delta_{n_1,n_2}\right)}\Re\left\{\left[\boldsymbol{Y}_m^{-1}\right]_{k_1,n_1}\left[\boldsymbol{Y}_m^{-1}\right]_{n_2,k_2} + \left[\boldsymbol{Y}_m^{-1}\right]_{k_1,n_2}\left[\boldsymbol{Y}_m^{-1}\right]_{n_1,k_2}\right\} \end{aligned} \tag{A.159}$$

$$\begin{aligned} &\frac{\partial^2 b}{\partial \left[\Re\{\boldsymbol{Y}_m\}\right]_{k_1,k_2} \partial \left[\Im\{\boldsymbol{Y}_m\}\right]_{n_1,n_2}} \\ &\quad = \frac{1}{t}\left(2-\delta_{k_1,k_2}\right)\Im\left\{\left[\boldsymbol{Y}_m^{-1}\right]_{k_1,n_2}\left[\boldsymbol{Y}_m^{-1}\right]_{n_1,k_2} - \left[\boldsymbol{Y}_m^{-1}\right]_{k_1,n_1}\left[\boldsymbol{Y}_m^{-1}\right]_{n_2,k_2}\right\} \end{aligned} \tag{A.160}$$

$$\begin{aligned} &\frac{\partial^2 b}{\partial \left[\Im\{\boldsymbol{Y}_m\}\right]_{k_1,k_2} \partial \left[\Im\{\boldsymbol{Y}_m\}\right]_{n_1,n_2}} \\ &\quad = \frac{2}{t}\Re\left\{\left[\boldsymbol{Y}_m^{-1}\right]_{k_1,n_1}\left[\boldsymbol{Y}_m^{-1}\right]_{n_2,k_2} - \left[\boldsymbol{Y}_m^{-1}\right]_{k_1,n_2}\left[\boldsymbol{Y}_m^{-1}\right]_{n_1,k_2}\right\}. \end{aligned} \tag{A.161}$$

The latter results show that $\boldsymbol{H}_{\boldsymbol{y}_m}$ can be computed with complexity $\mathcal{O}(M_T^4)$. Hence, the complexity to compute the Hessian (A.144) is $\mathcal{O}(M_T^2 N^2 + M_T^4 N)$.

To solve the KKT system (A.155), we formulate it as

$$\begin{aligned} \boldsymbol{H}\boldsymbol{\Delta v} + \boldsymbol{G}^T \boldsymbol{u} \quad &= -\boldsymbol{g} \\ \boldsymbol{G}\boldsymbol{\Delta v} \quad &= \boldsymbol{1}_{4N} \otimes \boldsymbol{b} - \boldsymbol{G}\boldsymbol{v}. \end{aligned} \tag{A.162}$$

The first equation of (A.162) gives

$$\boldsymbol{\Delta v} = -\boldsymbol{H}^{-1}\left(\boldsymbol{g} + \boldsymbol{G}^T \boldsymbol{u}\right). \tag{A.163}$$

Substituting (A.163) in the second equation of (A.162) yields

$$-\boldsymbol{G}\boldsymbol{H}^{-1}\boldsymbol{G}^T \boldsymbol{u} = \boldsymbol{G}\boldsymbol{H}^{-1}\boldsymbol{g} + \boldsymbol{1}_{4N} \otimes \boldsymbol{b} - \boldsymbol{G}\boldsymbol{v}. \tag{A.164}$$

Hence, we first solve (A.164). Subsequently, we substitute the result in (A.163).

To solve the KKT system in a computationally efficient way, we use that

$$\begin{aligned} \boldsymbol{G}\boldsymbol{H}^{-1}\boldsymbol{G}^T &= \boldsymbol{G}_{\boldsymbol{W}}\boldsymbol{H}_{\boldsymbol{W}}^{-1}\boldsymbol{G}_{\boldsymbol{W}}^T + \left(\boldsymbol{I}_{4N} \otimes \boldsymbol{G}_{\boldsymbol{y}}\right)\boldsymbol{H}_{\boldsymbol{y}}^{-1}\left(\boldsymbol{I}_{4N} \otimes \boldsymbol{G}_{\boldsymbol{y}}^T\right) \\ &\quad + \left(\boldsymbol{I}_{4N} \otimes \boldsymbol{G}_{\boldsymbol{z}}\right)\boldsymbol{H}_{\boldsymbol{z}}^{-1}\left(\boldsymbol{I}_{4N} \otimes \boldsymbol{G}_{\boldsymbol{z}}^T\right) \\ &= \boldsymbol{G}_{\boldsymbol{W}}\boldsymbol{H}_{\boldsymbol{W}}^{-1}\boldsymbol{G}_{\boldsymbol{W}}^T + \operatorname{diag}\left\{\boldsymbol{G}_{\boldsymbol{y}}\boldsymbol{H}_{\boldsymbol{y}_1}^{-1}\boldsymbol{G}_{\boldsymbol{y}}^T, \ldots, \boldsymbol{G}_{\boldsymbol{y}}\boldsymbol{H}_{\boldsymbol{y}_{4N}}^{-1}\boldsymbol{G}_{\boldsymbol{y}}^T\right\} \\ &\quad + \operatorname{diag}\left\{\boldsymbol{G}_{\boldsymbol{z}}\boldsymbol{H}_{\boldsymbol{z}_1}^{-1}\boldsymbol{G}_{\boldsymbol{z}}^T, \ldots, \boldsymbol{G}_{\boldsymbol{z}}\boldsymbol{H}_{\boldsymbol{z}_{4N}}^{-1}\boldsymbol{G}_{\boldsymbol{z}}^T\right\}. \end{aligned} \tag{A.165}$$

To compute (A.165), it can be exploited that each row of $\boldsymbol{G}_{\boldsymbol{W}}$ contains only $\mathcal{O}(M_T+N)$ non-zero entries. Also, the rows of $\boldsymbol{G}_{\boldsymbol{y}}$ and $\boldsymbol{G}_{\boldsymbol{z}}$ contain only $\mathcal{O}(M_T)$ non-zero entries. Using these results, it can be shown that the complexity for solving the KKT system is dominated by the inversion of $\boldsymbol{H}$, which is $\mathcal{O}(M_T^3 N^3 + M_T^6 N)$ due to the block structure of this matrix.

A.11 Proof of Proposition 5.2

If $\boldsymbol{w}$ satisfies the constraint in (5.148) with strict inequality, then a smaller value of the objective function can be obtained by scaling $\boldsymbol{w}$ with some positive factor that is strictly less than one. Hence, (5.149) holds.

Next, let us show that

$$|\boldsymbol{w}_q^H \hat{\boldsymbol{a}}_1(f_q + f_c)| \leq 1. \tag{A.166}$$

If $\boldsymbol{w}$ is an arbitrary weight vector with

$$|\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)| > 1, \tag{A.167}$$

then

$$\tilde{\boldsymbol{w}} = \frac{\boldsymbol{w}}{(\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c))^*} \tag{A.168}$$

satisfies

$$\tilde{\boldsymbol{w}}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \tilde{\boldsymbol{w}} < \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w}, \tag{A.169}$$

and

$$\left|\tilde{\boldsymbol{w}}^H \hat{\boldsymbol{a}}_1(f_q + f_c) - 1\right| + \epsilon_1(f_q + f_c) \, \|\tilde{\boldsymbol{w}}\|_2 < |\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c) - 1| + \epsilon_1(f_q + f_c) \|\boldsymbol{w}\|_2. \tag{A.170}$$

Therefore, the optimum weight vector has to satisfy (A.166).

To complete the proof, we show that $\boldsymbol{w}_q^H \hat{\boldsymbol{a}}_1(f_q + f_c)$ is real-valued and non-negative. If $\boldsymbol{w}$ is an arbitrary weight vector, which satisfies

$$|\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)| \leq 1 \tag{A.171}$$

and

$$\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c) \neq |\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)|, \tag{A.172}$$

then

$$\tilde{\boldsymbol{w}} = \frac{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)}{|\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)|} \boldsymbol{w} \tag{A.173}$$

satisfies

$$\tilde{\boldsymbol{w}}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \tilde{\boldsymbol{w}} = \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w}, \tag{A.174}$$

and

$$\left|\tilde{\boldsymbol{w}}^H \hat{\boldsymbol{a}}_1(f_q + f_c) - 1\right| + \epsilon_1(f_q + f_c) \, \|\tilde{\boldsymbol{w}}\|_2 < |\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c) - 1| + \epsilon_1(f_q + f_c) \|\boldsymbol{w}\|_2. \tag{A.175}$$

Due to (5.149), $\tilde{\boldsymbol{w}}$ and $\boldsymbol{w}$ cannot be optimum. This completes the proof.

A.12 Proof of Proposition 5.3

Due to (5.150) and (5.151), the solution of (5.148) does not change if we add the constraints

$$\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c) = \Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\right\}, \tag{A.176}$$

and

$$\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c) \leq 1. \tag{A.177}$$

This leads to the equivalent problem

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w} \\ \text{s.t.} \quad & \Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\right\} - \epsilon_1(f_q + f_c)\|\boldsymbol{w}\|_2 \geq 1 - \zeta(f_q) \\ & \Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\right\} \leq 1 \\ & \Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\right\} = \boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c). \end{aligned} \tag{A.178}$$

Let us relax (A.178) as

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w} \\ \text{s.t.} \quad & \Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\right\} - \epsilon_1(f_q + f_c)\|\boldsymbol{w}\|_2 \geq 1 - \zeta(f_q) \\ & \Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\right\} \leq 1. \end{aligned} \tag{A.179}$$

If $\boldsymbol{w}$ is a feasible point of (A.179) with

$$\Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\right\} \neq \boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c), \tag{A.180}$$

then

$$\tilde{\boldsymbol{w}} = \frac{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c) \Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\right\}}{|\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)|^2} \boldsymbol{w} \tag{A.181}$$

is another feasible point of (A.179), which leads to a strictly smaller value of the objective function. Consequently, $\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)$ has to be real-valued for the optimum weight vector of (A.179). The optimization problems (A.178) and (A.179) are therefore equivalent.

The constraint in (5.148) implies that

$$\|\boldsymbol{w}\|_2 \leq \frac{\zeta(f_q)}{\epsilon_1(f_q + f_c)}. \tag{A.182}$$

Hence, (A.179) is equivalent to the optimization problem

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w} \\ \text{s.t.} \quad & \Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\right\} - \epsilon_1(f_q + f_c)\|\boldsymbol{w}\|_2 \geq 1 - \zeta(f_q) \\ & \Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\right\} \leq 1 \\ & \|\boldsymbol{w}\|_2 \leq \tfrac{\zeta(f_q)}{\epsilon_1(f_q + f_c)}. \end{aligned} \tag{A.183}$$

To complete the proof, we show that (A.183) is equivalent to

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}_q} \boldsymbol{w} \\ \text{s.t.} \quad & \Re\left\{\boldsymbol{w}^H \hat{\boldsymbol{a}}_1(f_q + f_c)\right\} - \epsilon_1(f_q + f_c)\|\boldsymbol{w}\|_2 \geq 1 - \zeta(f_q) \\ & \|\boldsymbol{w}\|_2 \leq \tfrac{\zeta(f_q)}{\epsilon_1(f_q + f_c)}. \end{aligned} \tag{A.184}$$

Let $\breve{\boldsymbol{w}}_q$ denote the optimum weight vector of (A.184). Similar to the proof of Proposition 5.2, it can be shown that $\breve{\boldsymbol{w}}_q$ satisfies the constraint in the second line of (A.184) with equality. Hence,

$$\Re\left\{\breve{\boldsymbol{w}}_q^H \hat{\boldsymbol{a}}_1(f_q + f_c)\right\} = 1 - \zeta(f_q) + \epsilon_1(f_q + f_c)\|\breve{\boldsymbol{w}}_q\|_2. \tag{A.185}$$

The latter expression is less than or equal to one due to the constraint in the last line of (A.184). Consequently, $\breve{\boldsymbol{w}}_q$ is a feasible point of (A.183), i.e., the optimization problems (A.183) and (A.184) are equivalent.

Acronyms

1D	One-dimensional
2D	Two-dimensional
A/D	Analog-to-digital
CA	Covariance augmentation
CRB	Cramér-Rao bound
DOA	Direction-of-arrival
DFT	Discrete Fourier transform
FD	Fourier-domain
FDCA	Fourier-domain covariance augmentation
FIR	Finite impulse response
IDFT	Inverse discrete Fourier transform
INR	Interference-to-noise ratio
KKT	Karush-Kuhn-Tucker
LCMV	Linearly constrained minimum variance
LMI	Linear matrix inequality
MD	Multi-dimensional
ML	Maximum likelihood
MRA	Minimum redundancy array
MS	Manifold separation
MSE	Mean squared error
MUSIC	Multiple signal classification
MVDR	Minimum variance distortionless response
NMSE	Normalized mean squared error
NSD	Normalized signal distortion

NUA	Non-uniform array
OSINR	Output signal-to-interference-plus-noise ratio
RMSE	Root mean squared error
SDP	Semi-definite programming
SMI	Sample matrix inversion
SNR	Signal-to-noise ratio
SOCP	Second-order cone programming
SVD	Singular value decomposition
WFD	Weighted Fourier-domain
WLOG	Without loss of generality
UCA	Uniform circular array
ULA	Uniform linear array
URA	Uniform rectangular array

Frequently used symbols

Lower-case Latin letters:

$\boldsymbol{a}_l$	Steering vector of the lth source
$\underline{\boldsymbol{a}}_l$	Spatio-temporal steering vector of the lth source
c	Wave propagation speed
d_x, d_y, d_z	Baselines of a Cartesian grid in x, y, and z directions
$\boldsymbol{d}$	See (5.10)
$\boldsymbol{d}_{\text{cs}}$	See (5.71)
f	Baseband frequency
f_c	Carrier frequency
f_u	Maximum passband frequency of the source-of-interest
f_{AI}	Array interpolation null-spectrum function
f_{FD}	FD root-MUSIC null-spectrum function
f_{MS}	MS null-spectrum function
f_{MU}	MUSIC null-spectrum function
f_{WFD}	WFD root-MUSIC null-spectrum function
$\boldsymbol{n}$	Noise vector
$\underline{\boldsymbol{n}}$	Spatio-temporal noise vector
$p_{x_n}, p_{y_n}, p_{z_n}$	Coordinates of the nth sensor
s_l	Baseband waveform of the lth source
$\boldsymbol{s}$	Baseband source waveform vector
$\boldsymbol{t}$	Frequency response vector of the presteering stage
u_x, u_y, u_z	Electric angles
$u_{x,l}, u_{y,l}, u_{z,l}$	Electric angles of the lth source
$\boldsymbol{v}_{\boldsymbol{x},n}$	Eigenvector of $\boldsymbol{R}_{\boldsymbol{x}}$ corresponding to the nth largest eigenvalue λ_n
$\boldsymbol{w}$	Beamformer weight vector
$\underline{\boldsymbol{w}}$	Weight vector of FIR beamformers
$\boldsymbol{x}$	Snapshot vector
$\underline{\boldsymbol{x}}$	Spatio-temporal snapshot vector
$\underline{\boldsymbol{x}}_l$	Component of $\underline{\boldsymbol{x}}(kT_s)$ due to the lth source
y	Beamformer output
z	See (2.19)
z	See (2.31)

Capital Latin letters:

$\boldsymbol{A}$	Steering matrix
B_s	Bandwidth of the desired signal
$H(f, \boldsymbol{a})$	FIR beamformer frequency response for the baseband frequency f and the steering vector $\boldsymbol{a}$
$H_{1,\mathrm{des}}$	Desired FIR beamformer frequency response for the source-of-interest
K	Number of snapshots
L	Number of sources
M_T	FIR filter length
M_0	See (5.26)
N	Number of sensors
P_l	Received power of the lth source
P_{out}	Beamformer output power
$\boldsymbol{P}$	Source waveform covariance matrix
$\boldsymbol{R}_{\mathrm{i+n}}$	Interference-plus-noise covariance matrix
$\underline{\boldsymbol{R}}_{\mathrm{i+n}}$	Spatio-temporal interference-plus-noise covariance matrix
$\boldsymbol{R}_{\boldsymbol{x}}$	Array covariance matrix
$\underline{\boldsymbol{R}}_{\boldsymbol{x}}$	Spatio-temporal array covariance matrix
$\underline{\boldsymbol{R}}_l$	Component of $\underline{\boldsymbol{R}}_{\boldsymbol{x}}$ due to the lth source
$\underline{\boldsymbol{R}}_n$	Noise component of $\underline{\boldsymbol{R}}_{\boldsymbol{x}}$
S_l	Power spectral density of the lth source
$\mathcal{S}_{\epsilon_1}$	Set of presumed beamformer frequency responses for the desired signal
$\mathcal{S}_{\mathrm{lag}}$	Set of lags of the array geometry
T_{s}	Sampling time period
T_{se}	Beamformer sensitivity in the narrowband case
$\underline{T}_{\mathrm{se}}$	Beamformer sensitivity for FIR beamformers
$T_{p,n}$	Presteering delay of the nth sensor
$\bar{T}_p$	See (5.18)
$\boldsymbol{V}_{\boldsymbol{x},N}$	Noise subspace eigenvector matrix (2.11)
$\boldsymbol{V}_{\boldsymbol{x},S}$	Signal subspace eigenvector matrix (2.10)
$\boldsymbol{W}$	Weight matrix of FIR beamformers

Lower-case Greek letters:

α_w	Upper bound on the squared norm of the beamformer weight vector
$\boldsymbol{\delta}_l$	Estimation error of the steering vector of the lth source
ϵ_l	Upper bound on $\|\tilde{\boldsymbol{\delta}}_l\|_2$
$\breve{\epsilon}_l$	Upper bound on $\|\tilde{\boldsymbol{\delta}}_l\|_1$
θ	Elevation angle
θ_l	Elevation angle of the lth source
$\boldsymbol{\theta}$	Vector of source elevation angles
$\lambda_{\boldsymbol{x},n}$	nth largest eigenvalue of $\boldsymbol{R}_{\boldsymbol{x}}$
σ	Noise power
τ_n	Relative propagation delay of the desired signal to the nth sensor
φ	Azimuth angle
φ_l	Azimuth angle of the lth source
$\boldsymbol{\varphi}$	Vector of source azimuth angles

Capital Greek letters:

$\boldsymbol{\Delta}_{\boldsymbol{R}_{\boldsymbol{x}}}$	Array covariance matrix estimation error

Mathematical notation

Sets:

$\mathbb{R}$	Real numbers
$\mathbb{R}^{m\times n}$	Real $m \times n$ matrices
$\mathbb{C}$	Complex numbers
$\mathbb{C}^{m\times n}$	Complex $m \times n$ matrices
$L^2([-\pi,\pi])$	Lebesgue measurable functions with finite L^2-norm over the interval $[-\pi,\pi]$
$\mathcal{H}_+^{(M)}$	Positive semidefinite $M \times M$ matrices
$o(f)$	Functions with asymptotic growth rate less than that of f
$\mathcal{O}(f)$	Functions with asymptotic growth rate less than or equal to that of f
$\Theta(f)$	Functions with asymptotic growth rate equal to that of f

Vectors and matrices:

$\mathbf{0}$	Matrix of zeroes of conformable dimensions
$\mathbf{0}_m$	$m \times 1$ vector of zeroes
$\mathbf{0}_{m\times n}$	$m \times n$ matrix of zeroes
$\mathbf{1}$	Matrix of ones of conformable dimensions
$\mathbf{1}_m$	$m \times 1$ vector of ones
$\mathbf{1}_{m\times n}$	$m \times n$ matrix of ones
$\boldsymbol{e}_l$	lth standard basis column-vector of conformable dimension
$\underset{(N)}{\boldsymbol{e}_l}$	lth standard basis vector of dimension $N \times 1$
$\boldsymbol{I}_N$	$N \times N$ identity matrix
$(\cdot)^T$	Transpose
$(\cdot)^H$	Hermitian transpose
$(\cdot)^*$	Complex conjugate
$(\cdot)^\dagger$	Moore-Penrose pseudo-inverse
$\otimes$	Kronecker matrix product
$\odot$	Schur-Hadamard element-wise matrix product
$\mathrm{diag}\{\boldsymbol{x}\}$	Diagonal matrix with diagonal entries $x_1, x_2, \ldots$
$\mathrm{diag}\{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots\}$	Block diagonal matrix with diagonal blocks $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$

$\mathrm{vec}\{\cdot\}$	Vectorization of a matrix by stacking the columns on top of each other
$\underline{\mathrm{vec}}\{\cdot\}$	Vectorization for a square matrix, leaving out the entries on the main diagonal
$\mathrm{tr}\{\cdot\}$	Trace of a square matrix
$\mathcal{R}(\boldsymbol{A})$	Column-space of matrix $\boldsymbol{A}$
$\mathbf{\Pi}_{\boldsymbol{A}}$	Orthogonal projection matrix onto the column-space of $\boldsymbol{A}$
$\preceq$	Matrix inequality between Hermitian matrices

Scalar products and norms:

$\lvert\cdot\rvert$	Absolute value
$\lVert\cdot\rVert_2$	Euclidean vector norm or spectral matrix norm
$\lVert\cdot\rVert_1$	Vector 1-norm
$\lVert\cdot\rVert_\infty$	Vector ∞-norm
$\lVert\cdot\rVert_F$	Frobenius norm
$\lVert\cdot\rVert_{L^2}$	L^2-norm
$\langle\cdot,\cdot\rangle$	Scalar product for the Hilbert space $L^2([-\pi,\pi])$

Miscellaneous:

$\Re\{\cdot\}$	Real part
$\Im\{\cdot\}$	Imaginary part
$\mathrm{E}\{\cdot\}$	Statistical expectation
$\doteq$	Equality up to a scaling factor
$\approx$	Approximately equal
$\hat{}$	Estimate of
$\delta_{n,m}$	Kronecker delta
$\delta(\cdot)$	Dirac delta
$\forall$	For all
s.t.	Subject to
$F_f[m]$	mth Fourier series coefficient of f
$S_M\{f\}$	Fourier series expansion of f truncated to $2M+1$ terms

Bibliography

[Abr81] Y. I. Abramovich, “Controlled method for adaptive optimization of filters using the criterion of maximum SNR,” *Radio Eng. Electron. Physics*, vol. 26, pp. 87–95, 1981.

[AGGS98] Y. I. Abramovich, D. A. Gray, A. Y. Gorokhov, and N. K. Spencer, “Positive-definite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays. I. Fully augmentable arrays,” *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2458–2471, 1998.

[ASG99a] Y. I. Abramovich, N. K. Spencer, and A. Y. Gorokhov, “Positive-definite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays. II. Partially augmentable arrays,” *IEEE Trans. Signal Process.*, vol. 47, no. 6, pp. 1502–1521, 1999.

[ASG99b] ——, “Resolving manifold ambiguities in direction-of-arrival estimation for nonuniform linear antenna arrays,” *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2629–2643, 1999.

[ASG01] ——, “Detection-estimation of more uncorrelated Gaussian sources than sensors in nonuniform linear antenna arrays. I. Fully augmentable arrays,” *IEEE Trans. Signal Process.*, vol. 49, no. 5, pp. 959–971, 2001.

[ASG03a] ——, “Detection-estimation of more uncorrelated Gaussian sources than sensors in nonuniform linear antenna arrays. II. Partially augmentable arrays,” *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1492–1507, 2003.

[ASG03b] ——, “Detection-estimation of more uncorrelated Gaussian sources than sensors in nonuniform linear antenna arrays. III. Detection-estimation nonidentifiability,” *IEEE Trans. Signal Process.*, vol. 51, no. 10, pp. 2483–2494, 2003.

[BAMCM97] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, “A blind source separation technique using second-order statistics,” *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 434–444, 1997.

[Bar83] A. Barabell, “Improving the resolution performance of eigenstructure-based direction-finding algorithms,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, vol. 8, Boston, MA, USA, April 1983, pp. 336–339.

[Bel60] R. Bellman, *Introduction to Matrix Analysis.* McGraw-Hill, New York, NY, USA, 1960.

[BG86] K. Buckley and L. Griffiths, “An adaptive generalized sidelobe canceller with derivative constraints,” *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 311–319, 1986.

[BK80] G. Bienvenu and L. Kopp, “Adaptivity to background noise spatial coherence for high resolution passive methods,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, vol. 5, Denver, CO, USA, April 1980, pp. 307–310.

[Böh85] J. Böhme, “Source-parameter estimation by approximate maximum likelihood and nonlinear regression,” *IEEE Trans. Ocean. Eng.*, vol. 10, no. 3, pp. 206–212, 1985.

[Böh86] ——, “Separated estimation of wave parameters and spectral parameters by maximum likelihood,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, vol. 11, Tokyo, Japan, April 1986, pp. 2819–2822.

[Bra83] D. H. Brandwood, “A complex gradient operator and its application in adaptive array theory,” *IEE Proceedings F — Communications, Radar and Signal Processing*, vol. 130, no. 1, pp. 11–16, 1983.

[Bre78] J. Brewer, “Kronecker products and matrix calculus in system theory,” *IEEE Trans. Circuits Syst.*, vol. 25, no. 9, pp. 772–781, 1978.

[Bre79] ——, “Correction to ’Kronecker products and matrix calculus in system theory’,” *IEEE Trans. Circuits Syst.*, vol. 26, no. 5, p. 360, 1979.

[BRK07] F. Belloni, A. Richter, and V. Koivunen, “DOA estimation via manifold separation for arbitrary array structures,” *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4800–4810, 2007.

[BV04] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge University Press, Cambridge, UK, 2004.

[Cap69] J. Capon, “High-resolution frequency-wavenumber spectrum analysis,” *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.

[Car98] J.-F. Cardoso, “Blind signal separation: Statistical principles,” *Proc. IEEE*, vol. 86, no. 10, pp. 2009–2025, 1998.

[Che82] E. W. Cheney, *Introduction to Approximation Theory*, 2nd ed. Chelsea Publishing Company, New York, NY, USA, 1982.

[Com88] J. Compton, R. T., “The relationship between tapped delay-line and FFT processing in adaptive arrays,” *IEEE Trans. Antennas Propag.*, vol. 36, no. 1, pp. 15–26, 1988.

[Cor88] T. J. Cornwell, “A novel principle for optimization of the instantaneous Fourier plane coverage of correlation arrays,” *IEEE Trans. Antennas Propag.*, vol. 36, no. 8, pp. 1165–1167, 1988.

[CY92] L. Chang and C.-C. Yeh, “Performance of DMI and eigenspace-based beamformers,” *IEEE Trans. Antennas Propag.*, vol. 40, no. 11, pp. 1336–1347, 1992.

[CZO87] H. Cox, R. Zeskind, and M. Owen, “Robust adaptive beamforming,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 10, pp. 1365–1376, 1987.

[DD94a] M. A. Doron and E. Doron, “Wavefield modeling and array processing, Part I — Spatial sampling,” *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2549–2559, 1994.

[DD94b] ——, “Wavefield modeling and array processing, Part II — Algorithms,” *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2560–2570, 1994.

[DD94c] ——, “Wavefield modeling and array processing, Part III — Resolution capacity,” *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2571–2580, 1994.

[DLS02] T. N. Davidson, Z.-Q. Luo, and J. F. Sturm, “Linear matrix inequality formulation of spectral mask constraints with applications to FIR filter design,” *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2702–2715, 2002.

[Dum07] B. Dumitrescu, *Positive Trigonometric Polynomials and Signal Processing Applications.* Springer, Dordrecht, The Netherlands, 2007.

[EC83] M. Er and A. Cantoni, “Derivative constraints for broad-band element space antenna array processors,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 6, pp. 1378–1393, 1983.

[EKKG05a] A. El-Keyi, T. Kirubarajan, and A. B. Gershman, “Robust adaptive beamforming based on the Kalman filter,” *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3032–3041, 2005.

[EKKG05b] ——, “Wideband robust beamforming based on worst-case performance optimization,” in *Proc. IEEE Workshop on Statistical Signal Processing*, Bordeaux, France, July 2005, pp. 265–270.

[EKKG06] ——, “Adaptive wideband beamforming with robustness against presteering errors,” in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Waltham, MA, USA, July 2006, pp. 11–15.

[Ell01] S. W. Ellingson, “Design and evaluation of a novel antenna array for azimuthal angle-of-arrival measurement,” *IEEE Trans. Antennas Propag.*, vol. 49, no. 6, pp. 971–979, 2001.

[FM94] L. B. Fertig and J. H. McClellan, “Dual forms for constrained adaptive filtering,” *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 11–23, 1994.

[Fri93] B. Friedlander, “The root-MUSIC algorithm for direction finding with interpolated arrays,” *Signal Process.*, vol. 30, no. 1, pp. 15–25, 1993.

[Fro72] I. Frost, O. L., “An algorithm for linearly constrained adaptive array processing,” *Proc. IEEE*, vol. 60, no. 8, pp. 926–935, 1972.

[FW92] B. Friedlander and A. J. Weiss, “Direction finding using spatial smoothing with interpolated arrays,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 2, pp. 574–587, 1992.

[FW93] ——, “Direction finding for wide-band signals using an interpolated array,” *IEEE Trans. Signal Process.*, vol. 41, no. 4, pp. 1618–1634, 1993.

[GB97] A. B. Gershman and J. F. Böhme, “A note on most favorable array geometries for DOA estimation and array interpolation,” *IEEE Signal Process. Lett.*, vol. 4, no. 8, pp. 232–235, 1997.

[GB08a] M. Grant and S. Boyd. (2008, Oct.) CVX: Matlab software for disciplined convex programming. [Online]. Available: http://stanford.edu/ boyd/cvx

[GB08b] ——, “Graph implementations for non-smooth convex programs,” in *Recent Advances in Learning and Control*, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer, London, UK, 2008, pp. 95–110.

[Ger99] A. B. Gershman, “Robust adaptive beamforming in sensor arrays,” *Int. J. Electronics and Communications*, vol. 53, pp. 305–314, 1999.

[GHNvD03] V. Genin, Y. Hachez, Y. Nesterov, and P. van Dooren, “Optimization problems over positive pseudopolynomial matrices,” *SIAM J. Matrix Anal. Appl.*, vol. 25, pp. 57–79, 2003.

[GM55] E. N. Gilbert and S. P. Morgan, “Optimum design of directive antenna arrays subject to random variations,” *Bell Syst. Tech. J.*, vol. 34, pp. 637–663, 1955.

[God97] L. C. Godara, “Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations,” *Proc. IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.

[GRP10] A. B. Gershman, M. Rübsamen, and M. Pesavento, “One- and two-dimensional direction-of-arrival estimation: An overview of search-free techniques,” *Signal Process.*, vol. 90, no. 5, pp. 1338–1349, 2010.

[GSS$^+$10] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, “Convex optimization-based beamforming,” *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 62–75, 2010.

[GvL96] G. H. Golub and C. H. van Loan, *Matrix Computations.* The Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[GW78] C. R. Greene and R. C. Wood, “Sparse array performance,” *J. Acoust. Soc. Amer.*, vol. 63, no. 6, p. 18661872, 1978.

[Hal03] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction.* Springer, New York, NY, USA, 2003.

[HJO04] P. Hyberg, M. Jansson, and B. Ottersten, “Array interpolation and bias reduction,” *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2711–2720, 2004.

[HJO05] ——, “Array interpolation and DOA MSE reduction,” *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4464–4471, 2005.

[HNS01] M. Hawkes, A. Nehorai, and P. Stoica, “Performance breakdown of subspace-based methods: Prediction and cure,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Salt Lake City, UT, USA, May 2001, pp. 4005–4008.

[HRKV92] S. Haykin, J. P. Reilly, V. Kezys, and E. Vertatschitsch, “Some aspects of array signal processing,” *IEE Proceedings F — Radar Signal Process.*, vol. 139, no. 1, pp. 1–26, 1992.

[Hud81] J. E. Hudson, *Adaptive Array Principles*, P. J. B. Clarricoats, G. Millington, E. D. R. Shearman, and J. R. Wait, Eds. Peter Peregrinus, Stevenage, UK, and New York, NY, USA, 1981.

[JD93] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, A. V. Oppenheim, Ed. Prentice Hall, Englewood Cliffs, NJ, USA, 1993.

[Kay93] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, A. V. Oppenheim, Ed. Prentice Hall, Upper Saddle River, NJ, USA, 1993.

[KB86a] M. Kaveh and A. Barabell, “Corrections to “The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise”,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 3, p. 633, 1986.

[KB86b] ——, “The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 2, pp. 331–341, 1986.

[LB05] R. G. Lorenz and S. P. Boyd, “Robust minimum variance beamforming,” *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, 2005.

[Lee56] J. Leech, “On the representation of $1, 2, \ldots, n$ by differences,” *J. London Math. Soc.*, vol. 31, pp. 160–169, 1956.

[LLGS10] L. Lei, J. P. Lie, A. B. Gershman, and C. M. S. See, “Robust adaptive beamforming in partly calibrated sparse sensor arrays,” *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1661–1667, 2010.

[LP95] J.-W. Liang and A. J. Paulraj, “On optimizing base station antenna array topology for coverage extension in cellular radio networks,” in *Proc. IEEE 45th Vehicular Technology Conf*, vol. 2, 1995, pp. 866–870.

[LST93] D. A. Linebarger, I. H. Sudborough, and I. G. Tollis, "Difference bases and sparse sensor arrays," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 716–721, 1993.

[LSW03] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, 2003.

[LSW04] ——, "Doubly constrained robust Capon beamformer," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2407–2423, 2004.

[MAK97] A. Manikas, A. Alexiou, and H. R. Karimi, "Comparison of the ultimate direction-finding capabilities of a number of planar array geometries," *IEE Proceedings — Radar, Sonar and Navigation*, vol. 144, no. 6, pp. 321–329, 1997.

[MD01] Y. Meurisse and J.-P. Delmas, "Bounds for sparse planar and volume arrays," *IEEE Trans. Inf. Theory*, vol. 47, no. 1, pp. 464–468, 2001.

[MM80] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays.* Wiley, New York, NY, USA, 1980.

[Mof68] A. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. Antennas Propag.*, vol. 16, no. 2, pp. 172–175, 1968.

[MP98] A. Manikas and C. Proukakis, "Modeling and estimation of ambiguities in linear arrays," *IEEE Trans. Signal Process.*, vol. 46, no. 8, pp. 2166–2179, 1998.

[MPL99] A. Manikas, C. Proukakis, and V. Lefkaditis, "Investigative study of planar array ambiguities based on "hyperhelical" parameterization," *IEEE Trans. Signal Process.*, vol. 47, no. 6, pp. 1532–1541, 1999.

[MSD01] A. Manikas, A. Sleiman, and I. Dacos, "Manifold studies of nonlinear antenna array geometries," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 497–506, 2001.

[MZ94] C. P. Mathews and M. D. Zoltowski, "Eigenstructure techniques for 2-D angle estimation with uniform circular arrays," *IEEE Trans. Signal Process.*, vol. 42, no. 9, pp. 2395–2407, 1994.

[Neu69] H. Neudecker, "Some theorems on matrix differentiation with special reference to Kronecker matrix products," *J. Amer. Stat. Assoc.*, vol. 64, pp. 953–963, 1969.

[NN94] Y. Nesterov and A. Nemirovsky, *Interior-Point Polynomial Methods in Convex Programming*, ser. Studies in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, USA, 1994, vol. 13.

[OS99] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 2nd ed., A. V. Oppenheim, Ed. Prentice Hall, Upper Saddle River, NJ, USA, 1999.

[PBNH85] S. U. Pillai, Y. Bar-Ness, and F. Haber, “A new approach to array geometry for improved spatial spectrum estimation,” *Proc. IEEE*, vol. 73, no. 10, pp. 1522–1524, 1985.

[PGW02] M. Pesavento, A. B. Gershman, and K. M. Wong, “Direction finding in partly calibrated sensor arrays composed of multiple subarrays,” *IEEE Trans. Signal Process.*, vol. 50, no. 9, pp. 2103–2115, 2002.

[PPL90] D. Pearson, S. U. Pillai, and Y. Lee, “An algorithm for near-optimal placement of sensor elements,” *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1280–1284, 1990.

[PRK86] A. Paulraj, R. Roy, and T. Kailath, “A subspace rotation approach to signal parameter estimation,” *Proc. IEEE*, vol. 74, no. 7, pp. 1044–1046, 1986.

[Pum93] H. C. Pumphrey, “Design of sparse arrays in one, two, and three dimensions,” *J. Acoust. Soc. Amer.*, vol. 93, no. 3, p. 16201628, 1993.

[REKGK10] M. Rübsamen, A. El-Keyi, A. B. Gershman, and T. Kirubarajan, “Robust broadband adaptive beamforming using convex optimization,” in *Convex Optimization in Signal Processing and Communications*, D. P. Palomar and Y. C. Eldar, Eds. Cambridge University Press, 2010, ch. 9, pp. 315–339.

[RGa] M. Rübsamen and A. B. Gershman, “Robust adaptive beamforming using multi-dimensional covariance fitting,” *IEEE Trans. Signal Process.*, submitted.

[RGb] ——, “Sparse array design for azimuthal direction-of-arrival estimation,” *IEEE Trans. Signal Process.*, submitted.

[RG08a] M. Rübsamen and A. B. Gershman, “Performance analysis of root-MUSIC-based direction-of-arrival estimation for arbitrary non-uniform arrays,” in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Darmstadt, Germany, March 2008, pp. 381–385.

[RG08b] ——, “Robust presteered broadband beamforming based on worst-case performance optimization,” in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Darmstadt, Germany, July 2008, pp. 340–344.

[RG08c] ——, “Root-MUSIC based direction-of-arrival estimation methods for arbitrary non-uniform arrays,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Las Vegas, NV, USA, March 2008, pp. 2317–2320.

[RG09a] ——, "Direction-of-arrival estimation for nonuniform sensor arrays: From manifold separation to Fourier domain MUSIC methods," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 588–599, 2009.

[RG09b] ——, "Search-free DOA estimation algorithms for nonuniform sensor arrays," in *Classical and Modern Direction-of-Arrival Estimation*, T. Engin Tuncer and B. Friedlander, Eds. Elsevier, 2009, ch. 5, pp. 161–184.

[RG10a] ——, "Robust adaptive beamforming based on multi-dimensional covariance fitting," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Dallas, TX, USA, March 2010, pp. 2538–2541.

[RG10b] ——, "Subspace-based direction-of-arrival estimation for more sources than sensors using planar arrays," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Ma'ale Hahamisha, Israel, Oct. 2010, pp. 21–24.

[RH89] B. D. Rao and K. V. S. Hari, "Performance analysis of root-MUSIC," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 1939–1949, 1989.

[RMB74] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 10, no. 6, pp. 853–863, 1974.

[RMZ99] J. Ramos, C. P. Mathews, and M. D. Zoltowski, "FCA-ESPRIT: A closed-form 2-D angle estimation algorithm for filled circular arrays with arbitrary sampling lattices," *IEEE Trans. Signal Process.*, vol. 47, no. 1, pp. 213–217, 1999.

[RPK86] R. Roy, A. Paulraj, and T. Kailath, "ESPRIT–A subspace rotation approach to estimation of parameters of cisoids in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 5, pp. 1340–1342, 1986.

[RR49] L. Rédei and A. Rényi, "On the representation of $1, 2, \ldots, N$ by differences," *Mat. Sb. (N.S.)*, vol. 24 (66), no. 3, pp. 385–389, 1949, (in Russian).

[Rud64] W. Rudin, *Principles of Mathematical Analysis.* McGraw-Hill, Tokyo, Japan, 1964.

[RV06] T. Roh and L. Vandenberghe, "Discrete transforms, semidefinite programming, and sum-of-squares representations of nonnegative polynomials," *SIAM J. Opt.*, vol. 16, pp. 939–964, 2006.

[SBL11] P. Stoica, P. Babu, and J. Li, "SPICE: A sparse covariance-based estimation method for array processing," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 629–638, 2011.

[Sch68] F. Schweppe, "Sensor-array data processing for multiple-signal sources," *IEEE Trans. Inf. Theory*, vol. 14, no. 2, pp. 294–305, 1968.

[Sch79] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," in *Proc. RADC Spectral Estimation Workshop*, Rome, NY, USA, Oct. 1979, pp. 243–258.

[SG98] D. V. Sidorovich and A. B. Gershman, "Two-dimensional wideband interpolated root-MUSIC applied to measured seismic data," *IEEE Trans. Signal Process.*, vol. 46, no. 8, pp. 2263–2267, 1998.

[SGLW03] S. Shahbazpanahi, A. B. Gershman, Z.-Q. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2257–2269, 2003.

[Sko01] M. Skolnik, *Introduction to Radar Systems.* McGraw-Hill, New York, NY, USA, 2001.

[Ste92] G. W. Stewart, "An updating algorithm for subspace tracking," *IEEE Trans. Signal Process.*, vol. 40, no. 6, pp. 1535–1541, 1992.

[Str97] P. Strobach, "Bi-iteration SVD subspace tracking algorithms," *IEEE Trans. Signal Process.*, vol. 45, no. 5, pp. 1222–1240, 1997.

[SWL03] P. Stoica, Z. Wang, and J. Li, "Robust Capon beamforming," *IEEE Signal Process. Lett.*, vol. 10, no. 6, pp. 172–175, 2003.

[TBI97] L. N. Trefethen and D. Bau III, *Numerical Linear Algebra.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1997.

[TCL93] I. Thng, A. Cantoni, and Y. H. Leung, "Derivative constrained optimum broad-band antenna arrays," *IEEE Trans. Signal Process.*, vol. 41, no. 7, pp. 2376–2388, 1993.

[TCL95] ——, "Constraints for maximally flat optimum broadband antenna arrays," *IEEE Trans. Signal Process.*, vol. 43, no. 6, pp. 1334–1347, 1995.

[TLT95] J. Tuthill, Y. H. Leung, and I. Thng, "Adaptive RLS filters with linear and quadratic constraints," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, vol. 2, Detroit, MI, USA, May 1995, pp. 1424–1427.

[TMS01] A. R. Thompson, J. M. Moran, and G. W. Swenson, *Interferometry and Synthesis in Radio Astronomy.* Wiley-Interscience, New York, 2001.

[Tse92] C.-Y. Tseng, "Minimum variance beamforming with phase-independent derivative constraints," *IEEE Trans. Antennas Propag.*, vol. 40, no. 3, pp. 285–294, 1992.

[TST95] J. K. Thomas, L. L. Scharf, and D. W. Tufts, "The probability of a subspace swap in the SVD," *IEEE Trans. Signal Process.*, vol. 43, no. 3, pp. 730–736, 1995.

[Tuy95] H. Tuy, “D.C. optimization: Theory, methods and algorithms,” in *Handbook of Global Optimization, Nonconvex Optimization and its Applications*, R. Horst and P. Pardalos, Eds. Kluwer Academic Publishers, Dordrecht, Netherlands, 1995, vol. 2, pp. 149–216.

[VGL03] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, “Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem,” *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, 2003.

[VGLM04] S. A. Vorobyov, A. B. Gershman, Z.-Q. Luo, and N. Ma, “Adaptive beamforming with joint robustness against mismatched signal steering vector and interference nonstationarity,” *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 108–111, 2004.

[vT02] H. L. van Trees, *Optimum Array Processing.* Wiley, New York, NY, USA, 2002.

[WBV97] S. P. Wu, S. Boyd, and L. Vandenberghe, “FIR filter design via spectral factorization and convex optimization,” in *Applied Computational Control, Signal and Communications*, B. N. Datta, Ed. Birkhäuser, 1997, pp. 215–245.

[WFS95] A. J. Weiss, B. Friedlander, and P. Stoica, “Direction-of-arrival estimation using MODE with interpolated arrays,” *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 296–300, 1995.

[WK83] M. Wax and T. Kailath, “Optimum localization of multiple sources by passive arrays,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 5, pp. 1210–1217, 1983.

[WLS$^+$04] Z. Wang, J. Li, P. Stoica, T. Nishida, and M. Sheplak, “Constant-beamwidth and constant-powerwidth wideband robust Capon beamformers for acoustic imaging,” *J. Acoust. Soc. Amer.*, vol. 116, no. 3, pp. 1621–1631, 2004.

[Yan95] B. Yang, “Projection approximation subspace tracking,” *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 95–107, 1995.

[ZHM96] M. D. Zoltowski, M. Haardt, and C. P. Mathews, “Closed-form 2-D angle estimation with rectangular arrays in element space or beamspace via unitary ESPRIT,” *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 316–328, 1996.

[ZM91] M. D. Zoltowski and C. P. Mathews, “Beamspace root-MUSIC for rectangular arrays, circular arrays, and nonredundant linear arrays,” in *Proc. IEEE Asilomar Conf. Signals, Systmes and Computers*, 1991, pp. 556–560.

[ZT02] S. Zhang and I. L. J. Thng, “Robust presteering derivative constraints for broadband antenna arrays,” *IEEE Trans. Signal Process.*, vol. 50, no. 1, pp. 1–10, 2002.

[ZW88] I. Ziskind and M. Wax, “Maximum likelihood localization of multiple sources by alternating projection,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 10, pp. 1553–1560, 1988.

Curriculum vitae

Personal

Name: Michael Rübsamen
Address: Roßdörfer Straße 48, 64287 Darmstadt, Germany
Phone: +49 151 41226256
E-Mail: michael.rubsamen@googlemail.com
Date of birth: March 6, 1980
Place of birth: Hachenburg, Germany

Education

10/2000–04/2006	RWTH Aachen, Germany Diplom-Ingenieur (with highest honors) in Electrical Engineering and Information Technology with the specialization Information- and Communication Engineering, final thesis on “Spatial statistics of wireless networks”, supervised by Professor Petri Mähönen and M.Sc. Janne Riihijärvi
04/2000–09/2000	RWTH Aachen, Germany Study of physics
1990–1999	Konrad-Adenauer Gymnasium Westerburg, Germany Abitur (university entrance certificate)

Work experience

Since 06/2006	Technische Universität Darmstadt, Germany Research scientist in the Communication Systems Group, focusing on array signal processing
09/2009–10/2009	Temasek Labs at NTU, Singapore Scientific collaboration on direction-of-arrival estimation
10/2004–07/2005	Bell-Labs, Crawford Hill, NJ, USA Research scholarship, working on maximum likelihood sequence estimators for fiber-optic communication systems with high spectral efficiency
02/2001–03/2001	Pfeiffer Vacuum AG, Asslar, Germany Internship in the research and development department

Awards

2010	Student best paper award (first prize) for “Subspace-based direction-of-arrival estimation for more sources than sensors using planar arrays” at the IEEE Sensor Array and Multichannel Signal Processing (SAM) Workshop, Ma’ale Hahamisha, Israel (co-authored with Professor Alex Gershman)
2008	Student paper contest finalist for “Root-MUSIC based direction-of-arrival estimation methods for arbitrary non-uniform arrays” at the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Las Vegas, NV, USA (co-authored with Professor Alex Gershman)
2003–2006	Scholarship from the German Academic Foundation (Studienstiftung des deutschen Volkes)
2003	Henry-Ford-II Student Award
2002	Philips Student Award

Teaching experience

Summer semesters 2008 and 2009	Teaching assistant for the lecture “Information theory II” at Technische Universität Darmstadt
Summer semester 2007	Teaching assistant for the lecture “Information theory I” at Technische Universität Darmstadt
Winter semesters 2006–2009	Teaching assistant for the “Communications Lab A” at Technische Universität Darmstadt

Others

1999–2000	Military service

Darmstadt, April 11, 2011

Erklärung laut §9 der Promotionsordnung

Ich versichere hiermit, dass ich die vorliegende Dissertation allein und nur unter Verwendung der angegebenen Literatur verfasst habe. Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Darmstadt, 11. April 2011,